Received: 20 December 2018 | Revised: 12 March 2019 | Accepted: 25 March 2019

DOI: 10.1002/jbio.201800478

FULL ARTICLE

Journal of BIOPHOTONICS

Fast and reliable determination of *Escherichia coli* susceptibility to antibiotics: Infrared microscopy in tandem with machine learning algorithms

Uraib Sharaha[1] | **Eladio Rodriguez-Diaz**[2,3] | **Orli Sagi**[4] | **Klaris Riesenberg**[5] | **Ahmad Salman**[6*] | **Irving J. Bigio**[2,7] | **Mahmoud Huleihel**[1*]

[1]Department of Microbiology, Immunology and Genetic, Ben-Gurion University of the Negev, Beer-Sheva, Israel

[2]Department of Biomedical Engineering, Boston University, Boston, Massachusetts

[3]Section of Gastroenterology, VA Boston Healthcare System, Boston, Massachusetts

[4]Microbiology Laboratory, Soroka University Medical Center, Beer-Sheva, Israel

[5]Infectious Diseases Department, Soroka University Medical Center, Beer-Sheva, Israel

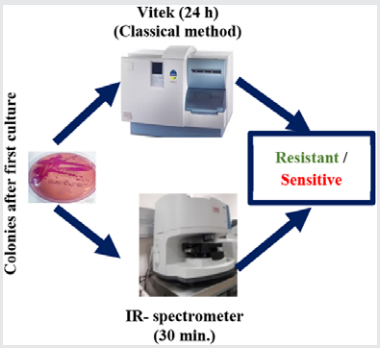
[6]Department of Physics, SCE—Shamoon College of Engineering, Beer-Sheva, Israel

[7]Department of Electrical & Computer Engineering, Boston University, Boston, Massachusetts

*Correspondence
Ahmad Salman
Department of Physics, SCE—Shamoon College of Engineering, Beer-Sheva 84100, Israel.
Email: ahmad@sce.ac.il
Prof. Mahmoud Huleihel
Department of Microbiology, Immunology and Genetics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel.
Email: mahmoudh@bgu.ac.il

Funding information
Shamoon College of Engineering (SCE)

Abstract



Antimicrobial drugs have an important role in controlling bacterial infectious diseases. However, the increasing resistance of bacteria to antibiotics has become a global health care problem. Rapid determination of antimicrobial susceptibility of clinical isolates is often crucial for the optimal antimicrobial therapy. The conventional methods used in medical centers for susceptibility testing are time-consuming (>2 days). Two bacterial culture steps are needed, the first is used to grow the bacteria from urine on agar plates to determine the species of the bacteria (~24 hours). The second culture is used to determine the susceptibility by growing colonies from the first culture for another 24 hours. Here, the main goal is to examine the potential of infrared microscopy combined with multivariate analysis, to reduce the time it takes to identify *Escherichia coli* susceptibility to antibiotics and to determine the optimum choice of antibiotic to which the bacteria will respond. *E coli* colonies of the first culture from patients with urinary tract infections (UTI) were examined for the bacterial susceptibility using Fourier-transform infrared (FTIR). Our results show that it is possible to determine the optimum choice of antibiotic with better than 89% sensitivity, in the time span of few minutes, following the first culture.

KEYWORDS
bacterial resistance to antibiotics, infrared spectroscopy, multivariate analysis, SVM

1 | INTRODUCTION

Although most of *Escherichia coli* strains are known as harmless bacteria that are commonly found in the lower intestine of endotherms, many *E coli* strains have been

Ahmad Salman, Irving J. Bigio and Mahmoud Huleihel contributed equally to this study.

implicated in a broad range of human diseases that extend from the gastrointestinal tract to extra-intestinal sites such as the urinary tract, bloodstream and central nervous system. *E coli* pathotypes cause much morbidity and mortality worldwide [1, 2].

Urinary tract infections (UTIs) are considered to be the most common bacterial infection, affecting 150 million people each year worldwide [3, 4], and the incidence is especially high in the places that suffer from lower socioeconomic status. UTIs occur four times more frequently in females than males [3]. Various studies have highlighted the importance of the medical and financial burdens associated with UTIs [5]; it was reported that the annual societal cost of these infections in the United States is approximately 3.5 billion US dollars [3]. Moreover, many patients with UTIs experience considerable impact related to quality of life [6].Treatment of UTIs is also necessary to reduce the risk of sequence to pyelonephritis [7]. Uropathogenic *Escherichia coli* (UPEC) is considered as the most common causative agent for both uncomplicated and complicated UTIs, since it constitutes 75% of these infections [8–12]. UPEC isolates have multiple implicated factors that promote colonization of the bacteria and infection in the urinary tract, such as toxins, fimbrial, adhesins and capsular polysaccharides [13, 14]. Multiple resistance to antimicrobial agents is increasing quickly in *E coli* isolates, leading to serious public health consequences, including increased mortality and morbidity, and complications associated with the therapeutic strategies for UTI [15, 16]. Uncontrolled use of antibiotics is considered to be the leading factor in the rise of multidrug resistance (MDR) in UPEC isolates [17]. UTI disease causes several complex symptoms, which induce physicians to begin experiential antibiotic treatments prior to obtaining culture results because urine culture susceptibility results take at least 48 hours [11, 18, 19]. Several studies reported that the prevalence of MDR *E coli* causing UTIs is increasing in Saudi Arabia [20], China [21], Japan [22] and the United States [23]. Therefore, reducing the time for susceptibility testing would lead to more timely determination of appropriate therapy and would reduce the emergence of antibiotic resistance.

There are several methods for testing bacterial susceptibility available today. These include the disk-diffusion method, the minimum inhibitory concentration (MIC) E-test and genotypic methods such as polymerase chain reaction and DNA hybridization methods [24]. The disk-diffusion test is the most common method used widely in clinical bacteriology labs. Although this method has advantages, like low cost and efficiency, a significant disadvantage of this method is that it is slow (requiring ~48 hours to obtain results) [24]. The genotypic methods are expensive and are not practical for routine clinical use [25].

Fourier-transform infrared (FTIR) microscopy is a powerful tool frequently used for chemical analysis, given its potential to provide detailed information on the chemical composition of materials at the molecular level [26]. It has been used in various disciplines, including material science, biochemistry and biomedical science due to the benefits of high sensitivity (SE), rapidity, simplicity and low cost [26, 27]. This technique has been proven to be sensitive for the identification of cancer cells and cancers [28–33], stem cells [34], virally infected cells [35, 36] and microorganisms [37].

In the field of bacteriology, FTIR spectroscopy has been successfully applied for detection, identification and classification of bacteria belonging to different species, particularly foodborne pathogens such as Listeria [38], *E coli* and Salmonella [39], Staphylococcus [40], Yersinia [41] and Bacillus [42]. Additionally, FTIR spectroscopy has a potential as a diagnostic tool in agricultural research, and has been used for classification, identification and discrimination of fungal isolates [43–45]. These studies have clearly indicated that FTIR microcopy, combined with multivariate analysis methods, has great potential for the identification and detection of biological samples [31]. Recently, we reported on the use of FTIR microscopy for assessing the susceptibility of bacteria to several antibiotics [46, 47]. In another previous study, FTIR-ATR (attenuated total reflection) was also used for the classification of *E coli* susceptibility into sensitive and resistant to cephalothin, utilizing classification by artificial neural network [48].

In this research, we continued evaluating the potential of the FTIR microscopy technique in tandem with multivariate analyses for rapid detection and identification of resistance of *E coli* isolates to all the antibiotic that are commonly used in the Soroka University Medical Center (SUMC), some of these being new and not previously published: co-trimoxazole, piperacillin/tazobactam, fosfomycin, amikacin, meropenem and ertapenem. Furthermore, we include in this manuscript updated data of *E coli* isolates; 791 samples compared to 496 samples used in the previously published results for ampicillin, cefuroxime, ceftriaxone, ciprofloxacin, sulfamethoxa, trimeth, amoxicillin clavulA, gentamicin, ceftazidime, nitrofurantoin, nalidixic acid and ofloxacin. Increasing the number of samples shows that the earlier classification results are reproducible, and that the method is rigorous, providing another step toward demonstrating proof-of-concept of the translational potential of this technique.

2 | MATERIALS AND METHODS

2.1 | Bacterial samples collection

Bacterial samples of *E coli* derived from patients with UTIs, from midstream specimens of urine, were examined by FTIR

microscopy for their susceptibility to all the antibiotics commonly used to treat these infections at SUMC. This study was carried out under the approval of the Institutional Ethical Helsinki Board (reference number 0398-15-SOR). These antibiotics are ampicillin, cefuroxime, ceftriaxone, ciprofloxacin, sulfamethoxa, trimeth, amoxicillin clavulA, gentamicin, ceftazidime, nitrofurantoin, nalidixic acid, ofloxacin, cotrimoxazole, piperacillin/tazobactam, fosfomycin, amikacin, meropenem and Ertapenem.

A total of 791 urine samples were obtained from 791 different patients with *E coli* UTI and sent to the bacteriology laboratories in SUMC for identification of their species by matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF MS) methods and for determination of susceptibility to antibiotics by MIC. These samples were supplied to us as colonies on MacConkey agar plates (first growth) from the bacteriology laboratories in SUMC.

2.2 | Sample preparation for FTIR microscopy measurements

Samples were taken directly from bacterial colonies on the plates and placed on zinc selenide slides using a sterile bacteriological loop under sterile conditions, dried at room temperature for ~15 minutes and examined by FTIR microscopy.

2.3 | FTIR measurements

FTIR microscopy was performed using a liquid-nitrogen-cooled, mercury-cadmium-telluride (MCT) detector, coupled to an FTIR spectrometer (Nicolet-i10 Infrared Microscope). To achieve a high signal-to-noise ratio (SNR), 128 co-added scans were collected in each measurement in the 600 to 4000 cm^{-1} wavenumber region in the transmission mode, with a spectral resolution of 4 cm^{-1}. These acquisition parameters enable us to obtain high quality and reproducible spectra.

2.4 | Spectral preprocessing

Various manipulations of the spectra must be performed to enhance the spectral features, facilitate spectral interpretation and analysis, and enable comparison between spectra from cultures with different thicknesses [26]. The acquired spectra were manipulated as follows: (a) atmospheric compensation was applied to eliminate influences of CO_2 and air humidity. (b) The spectra were smoothed using Savitzky-Golay algorithm with 13 points. (c) The spectra were cut into two regions, 900 to 1800 and 2800 to 3010 cm^{-1}, to exclude the water region 3010 to 4000 cm^{-1} and the region 1800 to 2800 cm^{-1} where there are no vibrational modes in biomolecules. (d) The spectral segments were then baseline-corrected using the concave rubber-band method [49], with 64 consecutive points and five iterations. (In using the concave rubber-band correction method, we divided the spectrum into 64 equal size ranges to construct the baseline.) (e) For absorbance spectra, the minimum ordinate-value of the ranges is determined. A polynomial function is fitted according to these minimum points. (f) Then we subtract this function from the spectrum to extract the baseline-corrected spectrum. This procedure was repeated for each iteration. (g) The corrected spectra were then normalized using vector normalization followed by offset correction. Using vector normalization, the average intensity in the spectral range is calculated and subtracted from the spectrum itself, and then the resulting spectrum is treated as vector that is normalized to unity.

Normalization of spectra eliminates the path-length variation [45, 50] and is a prerequisite to facilitate advanced statistical analysis of bacterial spectra [51]. All spectral manipulations were performed using the standard tools of (OPUS software version 7, Bruker, Germany).

2.5 | Statistical analysis

The multidimensional classification was performed using linear support vector machines (SVM) [52, 53]. Leave-one-out cross-validation (LOO) was used to estimate the success rate. Using LOO, all the spectra except one are used to train the system; the remaining spectrum's type is classified and compared to the known category (based on standard disc-culture methods). The LOO procedure is repeated for every individual sample spectrum, to estimate the statistical accuracy. It is important to mention that each spectrum included in the database represents a different isolate, which was obtained from different patient; thus leave-one-spectrum-out means leave-one-isolate-out.

The sensitive bacteria were defined as the "positive" state while the resistant bacteria were defined as the "negative" state in our analysis. The performances of the preliminary tests of the developed classifier were described using the following medical statistical terms: Specificity (SP) is defined as the probability to determine a known resistant isolate (using MIC) as resistant (by FTIR); and SE is defined as the probability to determine a sensitive isolate (using MIC) as sensitive (by FTIR). Accuracy (Acc) is defined as the combined probability to identify correctly resistant isolates as resistant and sensitive isolates as sensitive. Positive-predictive value (PPV) is defined as the probability that of the classifier correctly predicting an isolate to be sensitive to a specific antibiotic, while negative-predictive value (NPV) is defined as the probability of the classifier correctly determining an isolate to be resistant to a specific antibiotic.

2.6 | Ensemble analysis

Based on the fact that a number of antibiotics are available, and that the important issue is to help the physician in choosing one antibiotic that is most likely to work against the specific bacterial sample, an ensemble strategy of analysis was developed. The ensemble analysis is based on the combination results of a few antibiotics that have the best performance, and the main point is to choose one that has a high confidence level to treat the *E coli* isolate.

In the ensemble analysis, we ranked how sensitive a particular bacterial sample was to each antibiotic tested on it. This was accomplished by sorting in descending order the available antibiotics for a particular sample, based on the confidence levels obtained from each antibiotic's classifier. As a reminder, the confidence level output by each classifier provides a measure of how likely a sample is to be sensitive to a specific antibiotic. Thus, the higher the value of the confidence level, the higher the likelihood that the sample is sensitive to that particular antibiotic. Conversely, if, after ranking all available antibiotics, the highest ranking has a confidence level of 0.5 or less, it would mean that the sample is resistant to all those antibiotics.

In order to measure the performance of this type of analysis, we redefined SE and SP in this context. This is what is meant by "modified sensitivity and specificity." Ensemble SE is the probability of correctly identifying one (or more) "sensitive" antibiotic(s) from the available ones, based on the classical method gold standard. Ensemble SP is defined as the probability that the sample is determined to be resistant to all the available antibiotics that are also determined to be resistant by MIC.

3 | RESULTS

3.1 | Infrared absorption spectra of *E coli* bacteria

In this work, 791 isolates of UPEC were obtained from the microbiology lab, with all the information regarding their susceptibility to the different antibiotics as determined by MIC assay. These samples were analyzed using FTIR microscopy, combined with SVM classification to determine their susceptibility.

A representative infrared absorption spectrum of one of the *E coli* isolates after spectral manipulation is shown in Figure 1, for the high-wavenumber (2830-3000 cm^{-1}) and low-wavenumber (900-1790 cm^{-1}) regions.

The centroids of the major infrared absorption bands are labeled in the plot. These bands are attributed to the various vibrational modes of the different biomolecules that comprise the examined *E coli* bacterial samples. The absorption-bands centered at ~1079 cm^{-1} are mainly due to nucleic acids. Proteins are the main contributors to the absorption bands in the range 1590 to 1727 cm^{-1}; fatty acids are the main contributors to the absorption bands centered at 1402 cm^{-1}; and the absorption bands in the range 900 to 1200 cm^{-1} are mainly due to carbohydrates. Table 1

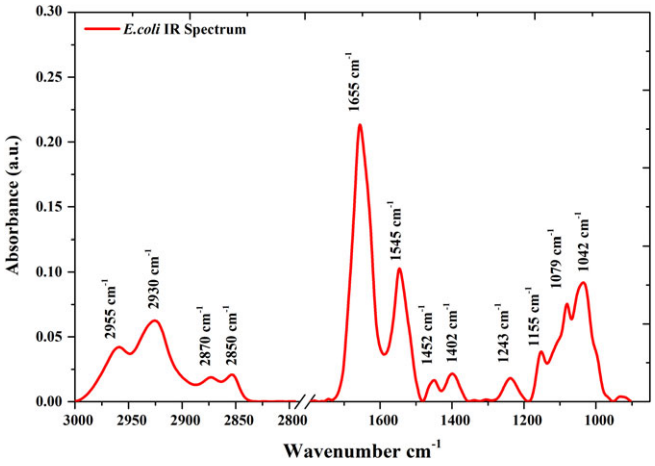


FIGURE 1 A representative FTIR absorption spectrum of an *Escherichia coli* isolate after manipulation in two regions: high wavenumber region (2830-3000 cm^{-1}) and low-wavenumber region (900-1800 cm^{-1})

TABLE 1 Assignments of the functional groups associated with major vibrational bands in mid IR spectra of bacteria

Wavenumber cm^{-1}	Molecular vibrations of functional groups and biomolecule contributor
2955	C-H asymmetric stretching of -CH3 in fatty acids
2930	C-H asymmetric stretching of >CH2 in fatty acids
2870	C-H symmetric stretching of -CH3 in fatty acids
2850	C-H symmetric stretching of >CH2 in fatty acids
1739	>C=O stretching of lipid esters
1590-1727	Amide I band components of proteins
1655	Amide I of α-helical structures of proteins
1637	Amide I of β-pleated sheet structures of proteins
1480-1590	Amide II band of proteins
1452	C-H deformation of >CH2 in lipids proteins
1402	C=O symmetric stretching of COO- group in amino acids, fatty acids
1310-1240	Amide III band components of proteins
1243	P=O asymmetric stretching of phosphodiesters in phospholipids
1200-900	C-O-C, C-O dominated by ring vibrations in various polysaccharides
1079	P=O symmetric stretching in DNA, RNA and phospholipids

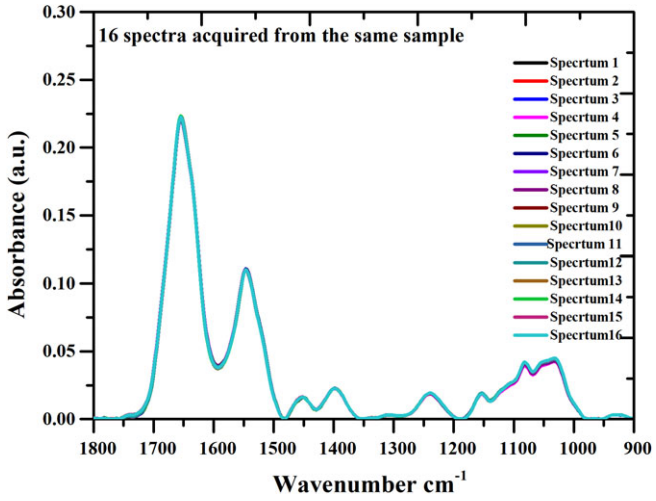


FIGURE 2 Sixteen IR microscopic absorbance spectra of one *Escherichia coli* isolate after preprocessing in the region 900 to 1800 cm^{-1}

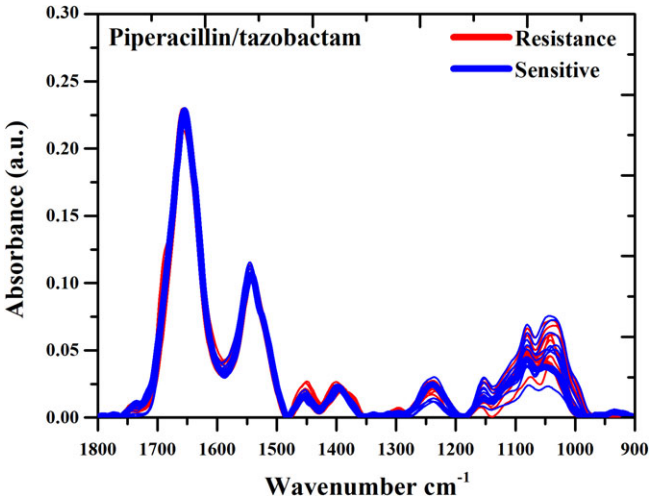


FIGURE 3 Infrared absorption spectra of *Escherichia coli* isolates sensitive (10 spectra) and resistant (10 spectra) to piperacillin/tazobactam in the region 900 to 1800 cm^{-1} after spectral manipulations

summarizes the functional groups associated with major vibrational bands in the mid-IR of bacteria that appear in Figure 1 [27, 37, 54–59].

3.2 | The reproducibility of the spectra

The differences among the spectra acquired from sensitive and resistant bacterial isolates are subtle; thus, it was important to have high a SNR and highly reproducible measurements in order to classify them with reasonable accuracy. For these reasons, the transmission mode was used to perform the measurements with 128 scans. At least 16 spectra from different sites on each sample were measured to examine the reproducibility of the spectra (Figure 2) and were averaged to increase the accuracy. The average of these spectra was used as the representative spectrum of each sample in the analysis. As can be seen from Figure 2, the spectra are almost overlaid, illustrating the high reproducibility of the measurements.

3.3 | FTIR spectra of different sensitive and resistant *E coli* isolates to the tested antibiotics

Figure 3 represents the spectra of 20 different isolates of *E coli* that were found to be variously sensitive or resistant to piperacillin/tazobactam. It can be seen that the spectra are similar, with significant overlap, but with a significant degree of variation of certain spectral components.

3.4 | Multivariate analysis

Advanced computational methods were used to analyze the pretreated spectra, and classify them based on the sample's susceptibility, as resistance or sensitive to a specific antibiotic. The classification was based on the small spectral differences resulting from the outcome of bacterial isolate's mutations, which are associated with the development of resistivity.

Given that these spectral changes associated with resistance to a specific antibiotic are subtle, statistical and multivariate analysis methods were used to develop a diagnostic algorithm based on multidimensional pattern-recognition/machine learning, enabling us to classify the spectra. As examples, in Figure 4, show the average spectra, in the low-wavenumber region, of the different *E coli* isolates, based on their susceptibility to different antibiotics: Cotrimoxazole (a), Piperacillin/tazobactam (c), Ceftriaxone (e) and Ceftazidime (g) are presented. The resulting receiver operating characteristic (ROC) [60] curves for these antibiotics are shown respectively in Figures 4b,d,f,h.

The ROC curves demonstrate graphically the accuracy of the tests (the probability for correct determination of the tested samples as resistant or sensitive), quantitatively represented by the area under the curve (AUC) of the ROC plot.

Due to the high similarity among the spectra of the different *E coli* isolates regardless to their susceptibility to antibiotics, our classification problem is binary, by which the spectra should be differentiated into resistant or sensitive spectra to a specific antibiotic. For optimization of the classifier parameters, the LOO validation approach was used.

The classification results of the developed classifiers for the examined antibiotics are summarized in Table 2.

All the classification results that appear in Table 2 were derived using different classifiers for each of the examined antibiotics; thus, each has its own performance statistics in the dataset. An important issue is the confidence level of the

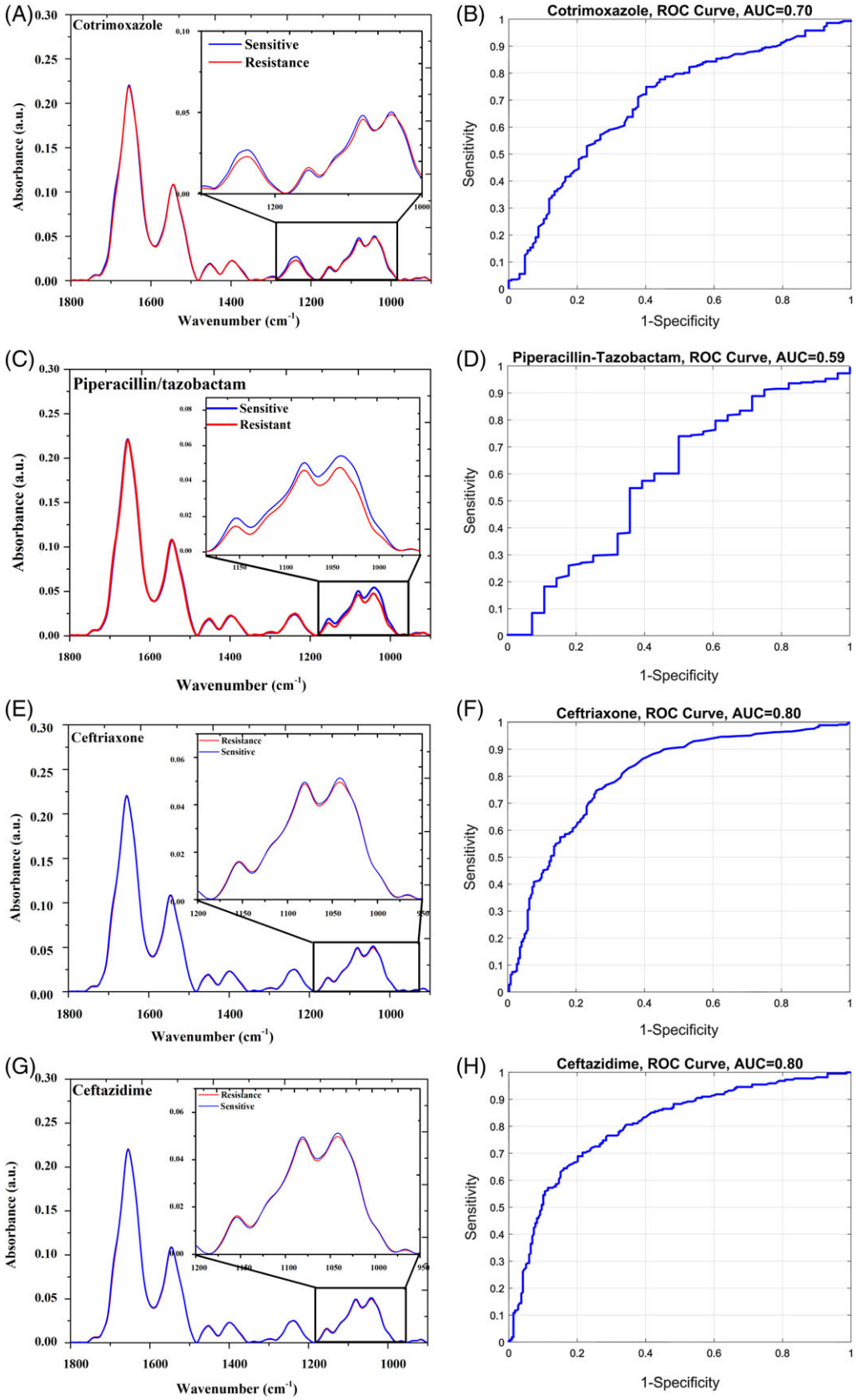


FIGURE 4 As exemplary for some of the investigated antibiotics, the average *Escherichia coli* IR spectra grouped based on the isolate susceptibility to (A) Cotrimoxazole, (C) Piperacillin/tazobactam, (E) Ceftriaxone and (G) Ceftazidime. The resulting ROC for classifying the bacterial samples based on their susceptibility to the same exemplary antibiotics respectively are shown in (B), (D), (F) and (H). The insets show detail of some of the spectral features (for *E coli* and for this antibiotic) that lead to the classification

TABLE 2 Classification performance for *Escherichia coli*

E coli	Resistant spectra	Sensitive spectra	SE	SP	Acc	PPV	AUC
Ampicillin	530	258	0.67	0.56	0.60	0.43	0.65
Cefuroxime	249	522	0.77	0.69	0.75	0.84	0.80
Ceftriaxone	221	470	0.76	0.71	0.75	0.85	0.80
Ciprofloxacin	219	482	0.83	0.73	0.80	0.87	0.84
Sulfamethoxa Trimeth	142	227	0.58	0.48	0.54	0.64	0.55
Amoxicillin ClavulA	119	529	0.70	0.55	0.67	0.87	0.66
Gentamicin	125	662	0.69	0.65	0.69	0.91	0.71
Ceftazidime	218	222	0.74	0.73	0.73	0.74	0.80
Nitrofurantoin	445	17	0.88	0.88	0.99	0.22	0.91
Nalidixic acid	73	34	0.81	0.82	0.91	0.67	0.89
Ofloxacin	80	25	0.87	0.92	0.97	0.70	0.94
Cotrimoxazole	269	514	0.71	0.61	0.68	0.81	0.68
Piperacillin/tazobactam	28	296	0.73	0.5	0.72	0.95	0.59

determination of an isolate as sensitive or resistant. The confidence level is related to the "distance" of the spectral feature values of the sample from the multidimensional class boundary.

Most of the misclassified isolates are near the boundary; thus, the confidence level of the classification for a specific sample may be high for one antibiotic and low for another. The ensemble analysis was carried out based on identification of an effective antibiotic, when N antibiotics exist.

As examples:

SE (1/1)—the probability of identifying the top-ranked antibiotic, when the top-ranked antibiotic is effective as determined by MIC. From a clinical perspective, in this case the performance achieved would be based on administering only the top-ranked antibiotic.

SE (1/2)—the probability of identifying at least one of the best two ranked antibiotics as effective when the top two are effective as determined by MIC. From a clinical perspective this case the performance achieved would be based on administering both top-ranked antibiotics.

SE (at least 1) measures the SE in correctly stating that the patient is sensitive to at least one antibiotic from the available ones. It does not, however, describe the performance of choosing a specific antibiotic. In a practical setting, this result lets you know whether your pool of available antibiotics has an antibiotic that can work or whether you need to look for other antibiotic.

Using this novel ensemble analysis, the individual classifiers were optimized for either Acc, or for PPV. The term "weighted" means that the classifier was optimized for either Acc or PPV, when the incidence of the SE and resistance are not equal and are weighted by the fraction of sensitive samples for a given antibiotic (a priori probability).

4 | DISCUSSION

FTIR microscopy is a rapid and powerful tool for the detections of small biochemical changes in cells and tissues [61]. Previous studies have used this technique successfully for taxonomic classification of different bacterial and fungal strains resulted from minor molecular changes [38–40, 42, 43, 62, 63]. Bacterial susceptibility to antibiotics may be manifested in minor molecular and biochemical changes in the bacterial cells [64]. Recently, Lechowicz et al, reported that FTIR microcopy combined with multivariate analysis methods has a great potential for classification UPEC strains' susceptibility to cephalothin [48]. In their study, they tested only one antibiotic and estimated their results for 109 samples by using the ATR/FTIR technique. Although they had high success rate, the number of samples used is deemed to be low, since multivariate analysis generally requires a large number of samples to yield reliable statistics.

MALDI-TOF mass spectrometry (MS) is one of the new methods, and many studies consider it to be an effective method for rapid testing of antibiotic susceptibility [65], and even for direct identification of bacteria causing extreme UTIs [66]. This technique depends on a database of known mechanisms of resistance; therefore, it is not powerful for the new, unknown mechanisms of resistance, and it cannot estimate the new resistant isolates. In addition, this technique may miss identification of mutant proteins with similar molecular weight to other cellular proteins.

In our previous publications, we reported that FTIR microscopy with multivariate analysis by SVM is a potential technique for rapid determination of UPEC susceptibility to several antibiotics [46, 47]. In this study, we have extended our research, examining the potential of this method for

TABLE 3 Ensemble classification performance for *Escherichia coli*

	SP	SE (at least 1)	SE (1/1)	SE (1/2)
Ensemble (Acc)	0.39	0.89	0.78	0.83
Weighted ensemble (Acc)	0.39	0.89	0.80	0.84
Ensemble (PPV)	0.59	0.83	0.69	0.75
Weighted ensemble (PPV)	0.59	0.83	0.74	0.78

successful determination of UPEC susceptibility to further antibiotics, to include all the antibiotics commonly used to treat these infections at SUMC. Some of these antibiotics are relatively new and are not prescribed frequently by physicians in clinics; thus, most of the *E coli* samples are sensitive to them: fosfomycin (662 sensitive and 6 resistance), amikacin (394 sensitive and 1 resistance), meropenem (419 sensitive and no resistance) and ertapenem (141 sensitive and no resistance). Unfortunately, it is a matter of time until *E coli* strains that are resistant to these new antibiotics will develop. Due to this situation, these antibiotics were excluded from our machine-learning analysis up to this point of time. We performed the analysis on two new antibiotics, cotrimoxazole and piperacillin/tazobactam, using 791 samples. For the other antibiotics (ampicillin, cefuroxime, ceftriaxone, ciprofloxacin, sulfamethoxa, trimeth, amoxicillin clavulA, gentamicin, ceftazidime, nitrofurantoin, nalidixic acid and ofloxacin) we performed new analysis to incorporate the larger number of the tested samples. The results revealed that when the number samples was increased from 495 samples to 791, the performance of our analysis method slightly decreased (Table 3) compared to our previous published results [46, 47]. In this study, a linear classifier was used to avoid overfitting. The classification results obtained using LOO are over-optimistic, since the training and testing datasets are statistically dependent. Testing a classification algorithm directly on naïve data (validated by classical tests) will require larger datasets, but increasingly sophisticated classifiers can be tested.

Thus, this study may be considered as an initial step toward providing proof-of-concept of the translational potential of this technique in clinical management of *E coli* bacterial infections in UTI.

5 | CONCLUSIONS

1. Infrared spectroscopy combined with machine-learning classification algorithms for pattern recognition is a powerful tool to determine the susceptibility of *E coli* bacteria to commonly used antibiotics.
2. The method can help physicians with a high confidence level to choose one or more antibiotics that are effective against the infecting *E coli* species, in time span of ~30 minutes, following the first culture.
3. Enlarging the database will enable using a nonlinear classifier that may improve the performance of the method.

ACKNOWLEDGMENT

Financial support by SCE internal research funding is gratefully acknowledged.

AUTHOR BIOGRAPHIES

Please see Supporting Information online.

REFERENCES

[1] M. A. Croxen, R. J. Law, R. Scholz, K. M. Keeney, M. Wlodarska, B. B. Finlay, *Clin. Microbiol. Rev.* **2013**, *26*, 822.
[2] J. B. Kaper, J. P. Nataro, H. L. Mobley, *Nat. Rev. Microbiol.* **2004**, *2*, 123.
[3] A. L. Flores-Mireles, J. N. Walker, M. Caparon, S. J. Hultgren, *Nat. Rev. Microbiol.* **2015**, *13*, 269.
[4] L. Perrault, S. Dahan, A. C. Iliza, J. LeLorier, G. G. Zhanel, *Can. J. Infect. Dis. Med. Microbiol.* **2017**, *2017*, 1.
[5] B. Foxman, R. Barlow, H. D'Arcy, B. Gillespie, J. D. Sobel, *Ann. Epidemiol.* **2000**, *10*, 509.
[6] M. Rosenberg, *Int. J. Antimicrob. Agents* **1999**, *11*, 247.
[7] S. Canbaz, Y. Peksen, A. T. Sunter, H. Leblebicioglu, M. Sunbul, *Int. J. Antimicrob. Agents* **2002**, *20*, 407.
[8] B. Foxman, *Infect. Dis. Clin. N. Am.* **2014**, *28*(1), 1.
[9] G. R. Nielubowicz, H. L. Mobley, *Nat. Rev. Urol.* **2010**, *7*, 430.
[10] G. Kahlmeter, *J. Antimicrob. Chemother.* **2003**, *51*, 69.
[11] K. Gupta, T. M. Hooton, K. G. Naber, B. Wullt, R. Colgan, L. G. Miller, G. J. Moran, L. E. Nicolle, R. Raz, A. J. Schaeffer, *Clin. Infect. Dis.* **2011**, *52*, e103.
[12] K. G. Naber, G. Schito, H. Botto, J. Palou, T. Mazzei, *Eur. Urol.* **2008**, *54*, 1164.
[13] R. Kucheria, P. Dasgupta, S. Sacks, M. Khan, N. Sheerin, *Postgrad. Med. J.* **2005**, *81*.
[14] S. Asadi, M. Kargar, K. Solhjoo, A. Najafi, S. Ghorbani-Dalini, *Jundishapur J. Microbiol.* **2014**, *7*(5):e9936. https://doi.org/10.5812/jjm.9936.
[15] O. J. Dyar, N. Q. Hoa, N. V. Trung, H. D. Phuc, M. Larsson, N. T. Chuc, C. S. Lundborg, *BMC Infect. Dis.* **2012**, *12*, 92.
[16] V. Niranjan, A. Malini, *Indian J. Med. Res.* **2014**, *139*, 945.
[17] T. D. Miles, W. McLaughlin, P. D. Brown, *BMC Vet. Res.* **2006**, *2*, 7.
[18] T. M. Hooton, *N. Engl. J. Med.* **2012**, *366*, 1028.
[19] D. S. Lo, H. H. Shieh, S. L. B. Ragazzi, V. H. K. Koch, M. B. Martinez, A. E. Gilio, *J. Bras. Nefrol.* **2013**, *35*, 93.
[20] A. M. Rafay, H. N. Nsanze, *Saudi Med. J.* **2003**, *24*, 261.
[21] P.-L. Ho, R. C. Wong, K.-S. Yip, S.-L. Loke, M. S. Leung, G. C. Mak, F. K. Chow, K. W. Tsang, T.-L. Que, C. S. Group, *Diagn. Microbiol. Infect. Dis.* **2007**, *59*, 439.

[22] K. Ishikawa, S. Hayakawa, S. Miyakawa, M. Kusaka, R. Shiroki, K. Hoshinaga, *J. Infect. Chemother.* **2005**, *11*, 44.
[23] D. F. Sahm, C. Thornsberry, D. C. Mayfield, M. E. Jones, J. A. Karlowsky, *Antimicrob. Agents Chemother.* **2001**, *45*, 1402.
[24] M. Balouiri, M. Sadiki, S. K. Ibnsouda, *J. Pharmaceut. Anal.* **2016**, *6*, 71.
[25] A. Davies, O. Billington, T. McHugh, D. Mitchison, S. Gillespie, *J. Clin. Microbiol.* **2000**, *38*, 3686.
[26] B. C. Smith, *Fundamentals of Fourier Transform Infrared Spectroscopy*, CRC Press, Boca Raton **2011**.
[27] M. J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H. J. Butler, K. M. Dorling, P. R. Fielden, S. W. Fogarty, N. J. Fullwood, K. A. Heys, *Nat. Protoc.* **2014**, *9*, 1771.
[28] P. D. Lewis, K. E. Lewis, R. Ghosal, S. Bayliss, A. J. Lloyd, J. Wills, R. Godfrey, P. Kloer, L. A. Mur, *BMC Cancer* **2010**, *10*, 640.
[29] M. J. Baker, E. Gazi, M. D. Brown, J. H. Shanks, P. Gardner, N. W. Clarke, *Br. J. Cancer* **2008**, *99*, 1859.
[30] M. A. Mackanos, *Trends Biotechnol.* **2010**, *28*, 317.
[31] G. Bellisola, C. Sorio, *Am. J. Cancer Res.* **2012**, *2*, 1.
[32] L. B. Mostaco-Guidolin, L. Bachmann, *Appl. Spectrosc. Rev.* **2011**, *46*, 388.
[33] C. Petter, N. Heigl, M. Rainer, R. Bakry, J. Pallua, G. Bonn, C. Huck, *Curr. Med. Chem.* **2009**, *16*, 318.
[34] J. G. Kelly, T. Nakamura, S. Kinoshita, N. J. Fullwood, F. L. Martin, *Analyst* **2010**, *135*, 3120.
[35] V. Erukhimovitch, M. Talyshinsky, Y. Souprun, M. Huleihel, *DNA Viruses: Methods and Protocols*, Humana Press, Totowa, NJ **2005**, pp. 161–172.
[36] A. Salman, V. Erukhimovitch, M. Talyshinsky, M. Huleihil, M. Huleihel, *Biopolymers* **2002**, *67*, 406.
[37] R. Davis, L. Mauer, *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, Vol. 2, Formatex Research Center, Badajoz, Spain **2010**, p. 1582.
[38] C. A. Rebuffo-Scheer, J. Schmitt, S. Scherer, *Appl. Environ. Microbiol.* **2007**, *73*, 1036.
[39] Y. Burgula, D. Khali, S. Kim, S. Krishnan, M. Cousin, J. Gore, B. Reuhs, L. Mauer, *J. Food Prot.* **2006**, *69*, 1777.
[40] H. Lamprell, G. Mazerolles, A. Kodjo, J. Chamba, Y. Noel, E. Beuvier, *Int. J. Food Microbiol.* **2006**, *108*, 125.
[41] G. E. Swann, S. Patwardhan, *Clim. Past* **2011**, *7*, 65.
[42] Z. Filip, S. Herrmann, J. Kubat, *Microbiol. Res.* **2004**, *159*, 257.
[43] A. Salman, A. Pomerantz, L. Tsror, I. Lapidot, R. Moreh, S. Mordechai, M. Huleihel, *Analyst* **2012**, *137*, 3558.
[44] A. Pomerantz, Y. Cohen, E. Shufan, Y. Ben-Naim, S. Mordechai, A. Salman, M. Huleihel, *J. Photochem. Photobiol. B Biol.* **2014**, *141*, 308.
[45] A. Salman, I. Lapidot, A. Pomerantz, L. Tsror, E. Shufan, R. Moreh, S. Mordechai, M. Huleihel, *BIOMEDO* **2012**, *17*, 0170021.
[46] A. Salman, U. Sharaha, E. Rodriguez-Diaz, E. Shufan, K. Riesenberg, I. J. Bigio, M. Huleihel, *Analyst* **2017**, *142*, 2136.
[47] U. Sharaha, E. Rodriguez-Diaz, K. Riesenberg, I. J. Bigio, M. Huleihel, A. Salman, *Anal. Chem.* **2017**, *89*, 8782.
[48] L. Lechowicz, M. Urbaniak, W. Adamus-Białek, W. Kaca, *Acta Biochim. Pol.* **2013**, 60, 713.
[49] A. Salman, E. Shufan, L. Zeiri, M. Huleihel, *Biochim. Biophys. Acta* **2013**, *1830*, 2720.
[50] C. Yu, J. Irudayaraj, *Biopolymers* **2005**, *77*, 368.
[51] C. Holt, D. Hirst, A. Sutherland, F. MacDonald, *Appl. Environ. Microbiol.* **1995**, *61*, 377.
[52] C. Cortes, V. Vapnik, *Mach. Learn.* **1995**, *20*, 273.
[53] V. N. Vapnik, V. Vapnik, *Statistical Learning Theory*, Wiley, New York **1998**.
[54] Z. Movasaghi, S. Rehman, D. I. ur Rehman, *Appl. Spectrosc. Rev.* **2008**, *43*, 134.
[55] L. J. Mauer, B. L. Reuhs *Wiley Handbook of Science and Technology for Homeland Security*, John Wiley and Sons inc, New Jersy **2008**, pp. 1–20.
[56] D. Naumann, *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Chichester, UK **2000**, pp. 102–131.
[57] K. Maquelin, C. Kirschner, L.-P. Choo-Smith, N. van den Braak, H. P. Endtz, D. Naumann, G. Puppels, *J. Microbiol. Methods* **2002**, *51*, 255.
[58] M. Beekes, P. Lasch, D. Naumann, *Vet. Microbiol.* **2007**, *123*, 305.
[59] B. R. Wood, M. A. Quinn, F. R. Burden, D. McNaughton, *Biospectroscopy* **1996**, *2*, 143.
[60] J. Platt, *Advances in Large Margin Classifiers* **1999**, 10, 61.
[61] F. L. Martin, J. G. Kelly, V. Llabjani, P. L. Martin-Hirsch, I. I. Patel, J. Trevisan, N. J. Fullwood, M. J. Walsh, *Nat. Protoc.* **2010**, *5*, 1748.
[62] C. A. Rebuffo, J. Schmitt, M. Wenning, F. von Stetten, S. Scherer, *Appl. Environ. Microbiol.* **2006**, *72*, 994.
[63] J. Kim, W. Schmid-Burgk, D. Claus, H. Kornhuber, *Arch. Psychiatr. Nervenkr.* **1982**, *232*, 299.
[64] G. S. Bbosa, N. Mwebaza, J. Odda, D. B. Kyegombe, M. Ntale, *Health* **2014**, *6*, 410.
[65] Y. Charretier, O. Dauwalder, C. Franceschi, E. Degout-Charmette, G. Zambardi, T. Cecchini, C. Bardet, X. Lacoux, P. Dufour, L. Veron, *Sci. Rep.* **2015**, *5*, 13944.
[66] X.-H. Wang, G. Zhang, Y.-Y. Fan, X. Yang, W.-J. Sui, X.-X. Lu, *J. Microbiol. Methods* **2013**, *92*, 231.

How to cite this article: Sharaha U, Rodriguez-Diaz E, Sagi O, et al. Fast and reliable determination of *Escherichia coli* susceptibility to antibiotics: Infrared microscopy in tandem with machine learning algorithms. *J. Biophotonics*. 2019;e201800478. https://doi.org/10.1002/jbio.201800478