2012

Pathogen identification with laser-induced breakdown spectroscopy: The effect of bacterial and biofluid specimen contamination

Q.I. Mohaidat

K. Sheikh

S. Palchaudhuri

Steven J. Rehse
University of Windsor

Pathogen identification with laser-induced breakdown spectroscopy: the effect of bacterial and biofluid specimen contamination

Qassem I. Mohaidat,[1] Khadija Sheikh,[2] Sunil Palchaudhuri,[3] and Steven J. Rehse[2,*]

[1]Department of Physics and Astronomy, Wayne State University, 666 W. Hancock, Detroit, Michigan 48201, USA

[2]Department of Physics, University of Windsor, 410 Sunset Avenue, Windsor, Ontario N9B 3P4, Canada

[3]Wayne State University, Department of Immunology and Microbiology, 7205 Scott Hall, Detroit, Michigan 48201, USA

*Corresponding author: rehse@uwindsor.ca

Received 23 September 2011; revised 20 December 2011; accepted 26 December 2011; posted 3 January 2012 (Doc. ID 154912); published 24 February 2012

In this paper, the potential use of laser-induced breakdown spectroscopy (LIBS) for the rapid discrimination and identification of bacterial pathogens in realistic clinical specimens is investigated. Specifically, the common problem of sample contamination was studied by creating mixed samples to investigate the effect that the presence of a second contaminant bacterium in the specimen had on the LIBS-based identification of the primary pathogen. Two closely related bacterial specimens, *Escherichia coli* strain ATCC 25922 and *Enterobacter cloacae* strain ATCC 13047, were mixed together in mixing fractions of 10:1, 100:1, and 1000:1. LIBS spectra from the three mixtures were reliably classified as the correct *E. coli* strain with 98.5% accuracy when all the mixtures were withheld from the training model and classified against spectra from pure specimens. To simulate a rapid test for the presence of urinary tract infection pathogens, LIBS spectra were obtained from specimens of *Staphylococcus epidermidis* obtained from distilled water and sterile urine. LIBS spectra from the urine-harvested bacteria were classified as *S. epidermidis* with 100% accuracy when classified using a model containing only spectra from other *Staphylococci* species and with 88.5% accuracy when a model containing five genera of bacteria was utilized. Bacterial specimens comprising five different genera and 13 classifiable taxonomic groups of species and strains were compiled in a library that was tested using external validation techniques. The importance of utilizing external validation techniques where the library is tested with data withheld from all previous testing and training of the model was revealed by comparing the results against "leave-one-out" cross-validation results. Last, the effect of using sequential models for the classification of a single unknown spectrum was investigated by comparing the misclassification of two closely related bacteria, *E. coli* and *E. cloacae*, when the classification was first performed using the five-genus bacterial library and then with a smaller model consisting only of *E. coli* and *E. cloacae* specimens. This result shows the utility of using successively more targeted analyses and models that use preliminary classifications from more general models as input.

OCIS codes: 140.3440, 170.4580, 280.1415, 300.6210, 300.6365.

1. Introduction

Only in the past decade has it begun to be recognized that laser-induced breakdown spectroscopy (LIBS) can be used for bacterial discrimination, classification, and identification [1–10]. The success of the technique has been due in part to several recent technological advances. One is the use of advanced computerized chemometric methods to analyze LIBS spectra [11], and another is the development of high-resolution, broadband echelle spectrometers

1559-128X/12/070B99-09$15.00/0

combined with sensitive CCD and intensified-CCD (ICCD) detectors [12]. Yet another is the availability of light-weight, inexpensive, high-powered lasers. By integrating these advances, LIBS can form the platform technology of an instrument that could be used for rapid point-of-care bacterial pathogen diagnosis. This offers a potentially faster, more reliable, and more robust diagnostic than other methods for the detection and identification of harmful pathogens in real time.

One example of a particular medical condition that could be greatly impacted by the use of a LIBS-based test is the diagnosis and treatment of urinary tract infections (UTIs). In 1995, 11.3 million women in the United States had at least one UTI, and 25%–40% of women in the United States aged 20–40 years have had a UTI [13]. Some estimate that UTIs cost at least $1 billion per year, and the overall cost of prescriptions alone to treat UTIs in 1995 was more than $218 million [13]. In hospitals, doctors order urine tests for their patients (most often women) to make sure that the kidneys and other organs are functioning well or when the patient may have an infection in his or her kidneys or bladder. Upon a positive diagnosis obtained from traditional culturing techniques (taking 24 to 48 h or from two to four clinical working days), therapy may be initiated with the proper antibiotic [14].

Staphylococcus epidermidis is a Gram-positive bacterium commonly found in the natural skin flora that sometimes causes human illness, including UTIs [15]. Infection caused by *S. epidermidis* is usually associated with medical devices such as indwelling catheters because it has the ability to form biofilms which will grow on those devices [16]. Hall and Snitzer investigated urinary tract infections in children without catheters that may be caused by *S. epidermidis*. In their study, they concluded that the presence of *S. epidermidis* bacteria in the urine culture should not automatically be considered a contaminant, especially when the clinical findings are compatible with urinary tract infection [17]. Therefore a reasonable UTI detection technology must be able to detect and identify a microorganism in a sample of urine (ordinarily sterile) and also should be unaffected by bacterial contaminations (small numbers of cells) from naturally occurring flora ("mixed samples"). This paper will investigate these two situations.

Rehse *et al.* have shown that LIBS has the capability of identifying the majority bacterium in mixed samples, specifically discriminating between *Mycobacterium smegmatis* and *Escherichia coli* [7]. In that study, bacteria were successfully identified in several mixture concentrations, including an 80 : 20 mixture. In real-world clinical situations where the LIBS-based identification is needed, it is anticipated that the pathogen causing the infection will be the majority species in any specimen of urine, blood, etc. (i.e., present in a mixed sample at concentrations far greater than any contaminant bacterial species). Investigations of these types of realistic mixing conditions are reported in this paper for two similar bacterial specimens relevant to UTIs, *E. coli* and *Enterobacter cloacae*. We also report on the use of LIBS to identify *S. epidermidis* cells isolated from otherwise sterile urine samples without washing. This was done to investigate if the presence of soluble salts and other biochemicals in the sterile urine interfered with the LIBS spectral identification when the model training set was constructed from bacterial spectra obtained from cells in deionized water samples.

A LIBS spectral library model consisting of five genus classes composed of 13 uniquely identifiable species and strains was tested using data sets withheld from the construction and training of the model (external validation). The ability of a discriminant function analysis (DFA) to accurately identify bacteria from spectra not represented in the model (i.e., grown on new or unique nutrient media or harvested from urine samples) will be presented.

Last, the use of a smaller, more focused model subsequent to a classification made by a model containing a more general but larger library was investigated. An improvement in the accuracy of classification was noted when similar bacterial spectra were tested in a model containing only the closely related spectra. This suggests that questionable identifications or identifications made on bacteria known to be highly similar to another type of bacteria should be confirmed or retested using a different model with a library specifically chosen to accentuate the anticipated differences.

2. Experiment

A. Bacterial Preparation of Mixed Samples and Sterile Urine Specimens

M. smegmatis bacteria were prepared by us in the same manner described in our previous work [7], while *S. epidermidis*, *E. coli* strain ATCC 25922, and *E. cloacae* strain ATCC 13047 specimens were prepared in a similar manner using similar growth media in the clinical microbiology lab of the Detroit Medical Center University Laboratories by Dr. Robert Mitchell. The two bacterial species used in this mixing test were specifically selected due to their similarity to each other. *E. coli* and *E. cloacae* both belong to a large family of Gram-negative rods called the enterobacteriaceae. These bacteria are referred to as "enteric pathogens" and can cause diarrheal disease [18]. These bacteria exhibit similar phenology and were expected to possess highly similar LIBS spectra, complicating the accurate classification of their mixtures. As well, both can be expected to be found in the human colon and in contaminated water or sewage. Therefore their presence together as a mixture is realistic. However, in a person exhibiting symptoms of enteric infection, only one of these pathogens would be responsible and would be dominant in any subsequently obtained culture. This is the motivation for our chosen mixing fractions.

Two separate suspensions of *E. coli* ATCC 25922 and *E. cloacae* ATCC 13047 were prepared prior to mixing, again using deionized water. To be a reasonable test, the two suspensions must have the same bacterial concentration. To do this, the turbidity or the optical density of each suspension was measured using a spectrophotometer at 600 nm (OD600). In this device, light is scattered as it passes through a bacterial suspension, and the amount of scatter is proportional to the number of bacteria in the suspension. The measured optical density was 0.78 for both. The measurements were conducted in the laboratory of Dr. Takeshi Sakamoto (Wayne State University, Department of Physics and Astronomy). After establishing the initial bacterial concentration, five separate mixtures were prepared with a ratio of *E. coli* ATCC 25922 (A) and *E. cloacae* ATCC 13047 (B) given by A_{1-x}: B_x with $x = 1.0$ 0.1, 0.01, 0.001 and 0.0 (three mixtures and two pure specimens). The three mixtures were prepared to give essentially 10∶1, 100∶1, and 1000∶1 titers, and the mixtures will be referred to by these mixing fractions in this paper. These are clinically realistic contamination mixing fractions. Multiple 1.5 ml aliquots of these mixtures were prepared, thoroughly agitated via vortex mixing, then centrifuged at 5000 rpm for 3 min to form a dense bacterial pellet. The supernatant was discarded, and 10 μl of the dense pellet was mounted on an agar surface prior to LIBS testing as has been described in our previous work [7].

To prepare urine specimens, *S. epidermidis* was collected from the growth plates and then suspended in separate 1.5 ml tubes of deionized water and sterile urine, using vortex mixing as described above to fully suspend the bacteria. After sitting in the tubes for some time (30 min or longer) samples of both were collected (using centrifugation and pipetting) without washing or any additional preparation to perform the LIBS experiment. All other control bacteria (*E. coli*, streptococci, and staphylococci) were collected from the growth plates of trypticase-soy agar and suspended in 1.5 ml deionized water prior to centrifugation and LIBS testing.

For the construction of the bacterial library, which contains LIBS spectra from almost every experiment conducted in our laboratory over the last three years, all bacteria were prepared in as similar a manner as possible given their preparation by numerous personnel over multiple years and in several different labs. The spectra therefore represent bacteria grown on hundreds of differently poured nutrition plates and mounted on dozens of nutrient-free agar substrates. Where noted, some of these bacteria have been intentionally exposed to various environmental conditions to test the alteration of the LIBS spectra (UV radiation exposure, starvation, autoclaving, etc.) as we have described in previous work.

B. LIBS Experimental Setup

The experimental setup used to perform LIBS on the bacterial samples is the same setup that has been described in our previous studies. 1064 nm pulses of 10 ns duration and approximately 10 mJ were used to ablate live bacteria mounted on nutrient-free bacto-agar plates. Specifically, 10 μl of bacterial pellet were transferred to 1.4% nutrient-free bacto-agar [7]. LIBS spectra were acquired in an argon environment at atmospheric pressure at a delay time of 2 μs after the ablation pulse, with an ICCD intensifier gate width of 20 μs duration. For the identification of bacteria in mixed samples and in sterile urine, five laser pulses were used to collect the spectra at one location, and five accumulations at five different locations were collected and averaged, resulting in a spectrum of 25 averaged laser pulses [7,9].

3. Results and Discussion

A. Identification of Bacteria in Mixed Samples

Figure 1 shows plots from a DFA of the LIBS spectra from pure samples of *M. smegmatis*, *E. coli* strain ATCC 25922, and *E. cloacae* strain ATCC 13047, as well as the mixtures with the mixing fractions mentioned earlier. In Fig. 1(a) all spectra were included in the construction of the model and were classified as unique classes. No relationship between the classes was provided. As can be seen from the plot of the first two discriminant function scores, the LIBS spectral fingerprint of the *E. cloacae* ATCC 13047 (group 1) bacterium was easily distinguishable from a spectrum obtained from a sample of pure *E. coli* ATCC 25922 (group 5). Spectra from the three mixtures (groups 2–4) were essentially indistinguishable from the spectra of *E. coli* ATCC 25922. When individual spectra from the mixtures were omitted from the model in a "leave-one-out" (LOO) cross validation they were always correctly classified as *E. coli* with 100% accuracy. This confirms the fact that the bacterial cells of *E. coli* ATCC 25922 were the dominant contributors to the LIBS spectrum in all the mixture samples.

In this analysis, notice that the discriminant function two (DF2) score of the centroid (which is the effective "center of mass" of the distribution of measurements) of the *M. smegmatis* samples was zero. On the other hand, the specimens of *E. coli* and *E. cloacae* (both pure and mixtures) possessed almost the same DF1 score (indicating their strong similarity). The interpretation of this is that the discrimination between the LIBS spectra from *E. coli* ATCC 25922 and *E. cloacae* ATCC 13047 and their mixtures was based on the difference between their DF2 scores only. There is a slight shift downward in the centroid of the 10∶1 mixture (downward toward the DFA space of the minority component) as expected. In our previous work we noted a similar shift in DFA space, and a corresponding decrease in accuracy, as the fraction of the majority bacterium dropped below 80%.

In Fig. 1(b) all of the mixture specimens were declassified so that the model was not constructed with mixtures and had never been tested or trained with a spectrum from a mixed sample. We will refer to this

type of testing of the model as "external validation," meaning the test spectra are external to the model, having been withheld from its construction. The unclassified mixtures [represented with the "*x*" symbols in Fig. 1(b)] were then classified using the model of pure *E. coli*, *E. cloacae*, and *M. smegmatis*. In this analysis, 67 of 68 spectra (98.5%) obtained from mixtures were classified correctly as *E. coli*. The remaining spectrum was classified incorrectly as *E. cloacae*. The success of this external validation of a model that was not created specifically to deal with mixtures demonstrates the strength of the DFA classification of the unknown specimens and again shows that a model need not contain every potential class that may be encountered to still be highly effective.

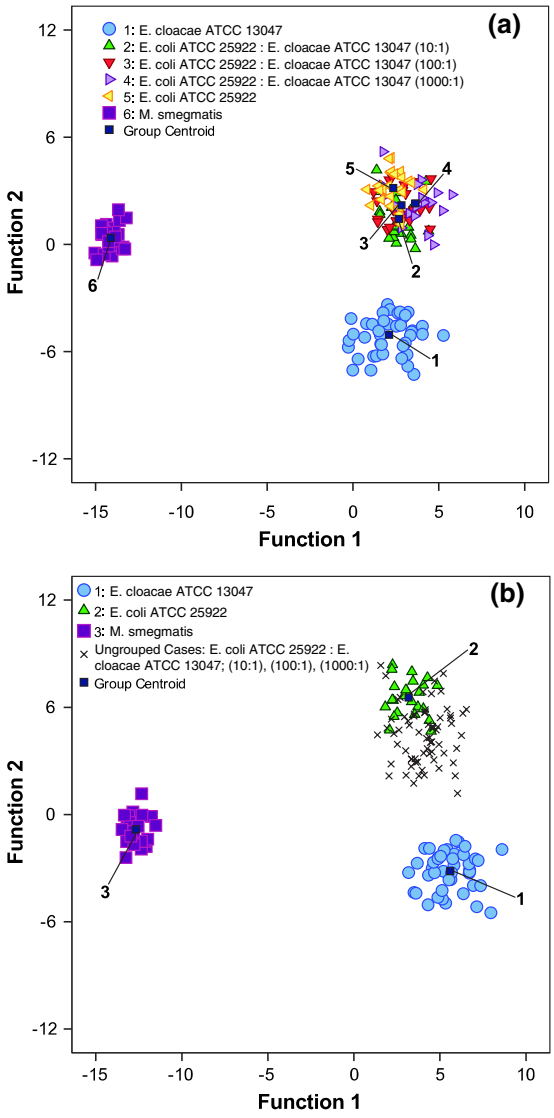

Fig. 1. (Color online) DFA plots showing the first two discriminant function scores from a DFA of LIBS spectra obtained from pure samples of three bacteria, *E. cloacae* strain ATCC 13047, *E. coli* strain ATCC 25922, and *M. smegmatis*, as well as three mixtures of the *E. coli* and *E. cloacae* at various mixing fractions. (a) All classes were input to the model, but no relationships were input, indicating the high similarity between the pure *E. coli* spectra and those of the mixtures. (b) Spectra from the three mixtures were declassified and tested against the model constructed only from spectra from pure specimens. All mixture spectra but one were correctly identified as *E. coli*.

These results suggest that spectra from *E. coli* ATCC 25922 bacteria could be identified with a high accuracy even in the presence of low concentrations of *E. cloacae* ATCC 13047, a common clinical contaminant. The result is in good agreement with what we obtained in a previous study [7] in which we observed that in clinical samples that may contain more than one microorganism the infection will most likely be caused by the microbe that has the higher concentration (often by orders of magnitude).

B. Identification of *S. epidermidis* Bacteria in a Sterile Urine Suspension

LIBS spectra from bacterial cells of *E. coli* strain C, *S. viridans*, and *S. epidermidis* that were suspended in water were acquired as described earlier. In addition to these three specimens, LIBS spectra were also collected from *S. epidermidis* bacterial cells that were suspended in sterile urine. Generally, normal urine consists of 96% water and 4% solutes. Organic solutes include urea, ammonia, creatinine, and uric acid. Inorganic solutes include sodium chloride, potassium sulfate, magnesium, and phosphorus. Figure 2 shows the results of a discriminant function analysis performed on the LIBS spectra of *S. epidermidis* cells harvested from the urine samples.

Figure 2(a) shows the first two discriminant function scores of a DFA performed on spectra acquired from the three specimens mentioned above. The urine-harvested samples were entered unclassified into this analysis, and none of the other bacterial cells used to construct this model had ever been exposed to urine or its constituent solutes at any time during growth. In Fig. 2(a), it can be seen that the LIBS spectral fingerprint from urine-exposed bacterial cells (unclassified "*x*" symbols) was essentially identical to the fingerprint from the same bacterial cells when harvested from water and was thus correctly classified with 100% accuracy. The effect of the presence of organic or inorganic solutes in the urine on the LIBS spectrum of the bacteria was negligible.

To verify our results in a more difficult test, the LIBS spectra from the *S. epidermidis* samples were tested against a model containing only staphylococci by adding *S. aureus* and *S. saprophyticus* to the model. This is shown in Fig. 2(b). Again, the LIBS spectra from *S. epidermidis* cells harvested from a urine suspension were entered as unclassified cases in a model composed of spectra from cells of the three staphylococci species obtained from water specimens. In this analysis all the LIBS spectra of *S. epidermidis* cells harvested from urine were again essentially identical to the spectra from bacterial cells that were collected from water and were again classified with 100% accuracy. When the spectra of *S. epidermidis* cells harvested from a urine specimen were unclassified in a

model containing every bacterial spectrum we have taken (discussed in Subsection 3.C), 88.5% of the spectra were correctly classified as *S. epidermidis* (in water).

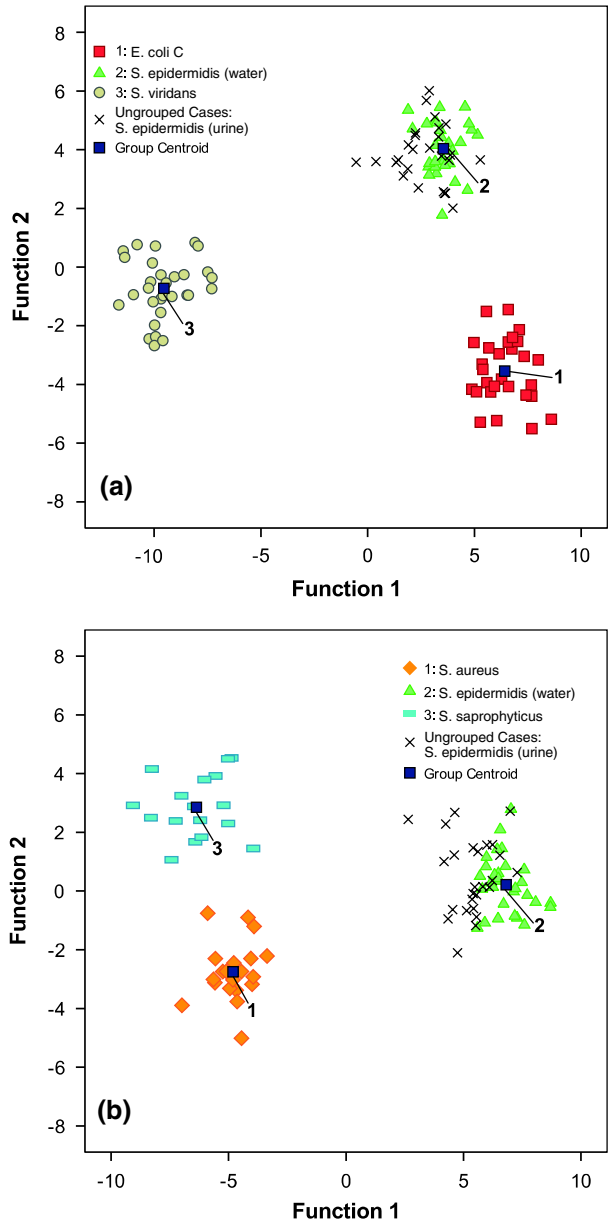

Fig. 2. (Color online) DFA plots showing the first two discriminant function scores from a DFA of LIBS spectra from *S. epidermidis* bacterial cells harvested from both urine and deionized water specimens. (a) The model was constructed from *E. coli* strain C, *S. epidermidis* cells harvested from water, and *S. viridans*. None of the spectra in the model were obtained from any urine-exposed cells. Spectra from *S. epidermidis* cells harvested from urine suspensions were tested with this model, and 100% of the specimens were correctly identified as *S. epidermidis*. (b) To complicate the test, the *S. viridans* and *E. coli* spectra were replaced with spectra from two other staphylococci species: *aureus* and *saprophyticus*. All cells were obtained from deionized water specimens. Spectra from *S. epidermidis* cells harvested from urine suspensions were tested with this model, and again 100% of the specimens were correctly identified as *S. epidermidis*.

C. Five-Genus Bacterial Spectral Library

LIBS spectral fingerprints collected from ongoing experiments over the past three years have been compiled to form a spectral library model. The library contains five bacterial genera composed of 13 unique bacterial constituents (strains and species) and 32 completely distinct sets of data (taken on different days, under different growth conditions, etc.) that can be omitted entirely from a DFA to allow a true external validation of the model. This is shown in Table 1. The library contains 669 LIBS spectra.

This library can be constructed, trained, and tested differently depending on whether the model consists of five classes (the five genera shown in column one of Table 1) or 13 classes (the 13 unique bacterial constituents shown in column two of Table 1). The resulting effect on classification, however, was not dramatic. This is shown in Fig. 3, which shows the first three discriminant function scores of a DFA of the 669 spectra when they were grouped into five classes [Fig. 3(a)] or 13 classes [Fig. 3(b)]. The first three DF scores are highly similar (but not identical) whether there were four discriminant functions calculated [Fig. 3(a)], or 12 [Fig. 3(b)]. To assist in a comparison of Fig. 3(a) and Fig. 3(b), the symbols of the 13 classes in Fig. 3(b) have been made uniform with the genus symbols in Fig. 3(a), and for clarity the group centroid and labels of the 13 classes are omitted. The striking similarity between Fig. 3(a) and Fig. 3(b) shows that the DFA does not need to "force" different bacterial spectra into arbitrary classes but that the LIBS spectra from the 13 bacterial taxonomic classifications naturally group in this way.

An external validation test was performed on the five-class, genus-level model of the bacterial spectra. To do this, one by one each of the 32 unique sets of data listed in Table 1 was withheld from the five-class model and entered as unclassified data. The DFA then assigned these spectra to one of the five genus classes. This was not possible with data where only one data set existed (i.e., set 15: *E. coli* Hfr-K-12 or set 17: *S. saprophyticus*), so those data were not classified in this way. Truth tables were constructed for each of the genera, and the results are shown in Table 2. For comparison, truth tables for a LOO cross validation are also shown. The truth tables compare rates of sensitivity (true positives) and specificity (true negatives). As expected, the sensitivity and selectivity are both (artificially) higher in the LOO cross validation. The truth table values shown for the external validation data are much more likely to represent the true sensitivity and specificity of this LIBS-based test. Sensitivities with our current apparatus range from 62% to 90%, and specificities are all greater than 95%. The staphylococci data possessed the lowest sensitivity, but these data represent the oldest data in the library, obtained quite some time ago before our testing, harvesting, and growth protocols were standardized. Nonetheless, we include them for completeness here. We expect

Table 1. Identities of the Classes Used to Construct a Spectral Library Model Composed of 669 Bacterial LIBS Spectra

Genus	Bacterial ID	Data set
1: *Escherichia*	1: *E. coli* ATCC 25922	1: *E. coli* ATCC 25922
	1: *E. coli* ATCC 25922	2: *E. coli* ATCC 25922/*E. cloacae* (10:1)
	1: *E. coli* ATCC 25922	3: *E. coli* ATCC 25922/*E. cloacae* (100:1)
	1: *E. coli* ATCC 25922	4: *E. coli* ATCC 25922/*E. cloacae* (1000:1)
	2: *E. coli* O157:H7 (EHEC)	5: *E. coli* O157:H7
	3: *E. coli* C	6: *E. coli* C
	3: *E. coli* C	7: *E. coli* C—cultured on MacConkey agar
	3: *E. coli* C	8: *E. coli* C—starved for 1 day
	3: *E. coli* C	9: *E. coli* C—starved for 4 days
	3: *E. coli* C	10: *E. coli* C—starved for 6 days
	3: *E. coli* C	11: *E. coli* C—starved for 8 days
	3: *E. coli* C	12: *E. coli* C—autoclaved
	3: *E. coli* C	13: *E. coli* C—UV exposed/killed
	4: *E. coli* HF4714	14: *E. coli* HF4714
	5: *E. coli* Hfr-K12	15: *E. coli* Hfr-K12
2: *Enterobacter*	6: *E. cloacae* ATCC 13047	16: *E. cloacae* ATCC 13047
3: *Staphylococcus*	7: *S. saprophyticus*	17: *S. saprophyticus*
	8: *S. aureus*	18: *S. aureus*
4: *Streptococcus*	9: *S. mutans*	19: *S. mutans*
	10: *S. viridans*	20: *S. viridans*
	10: *S. viridans*	21: *S. viridans*—starved for 1 day
	10: *S. viridans*	22: *S. viridans*—starved for 6 days
	10: *S. viridans*	23: *S. viridans*—starved for 9 days
	10: *S. viridans*	24: *S. viridans*—UV exposed/killed
	10: *S. viridans*	25: *S. viridans*—autoclaved
5: *Mycobacterium*	11: *M. smegmatis* WT	26: *M. smegmatis* WT—90% dilution
	11: *M. smegmatis* WT	27: *M. smegmatis* WT—60% dilution
	11: *M. smegmatis* WT	28: *M. smegmatis* WT—50% dilution
	11: *M. smegmatis* WT	29: *M. smegmatis* WT
	11: *M. smegmatis* WT	30: *M. smegmatis* WT—100% concentration
	12: *M. smegmatis* TE	31: *M. smegmatis* TE
	13: *M. smegmatis* TA	32: *M. smegmatis* TA

a more complete library, where all specimens are prepared with the same protocol, to yield sensitivities greater than 90%.

These sensitivities can be compared with alternate techniques. In a study on the use of real-time polymerase chain reaction (PCR) for the detection of methicillin-resistant *Staphylococcus aureus* (MRSA) in a hospital, a sensitivity of 96% was determined for the real-time PCR assay [19]. The gold-standard method of culture and count has a sensitivity of 100%. A more useful and meaningful metric is the positive predictive value (PPV). While the sensitivity is determined by taking the ratio of true positive test results to the actual number of positive specimens, the PPV is the ratio of true positive test results to the sum of the number of true positive and false positive test results. Because all "positive" results must be responded to with appropriate measures in a clinical setting (i.e., pharmaceutical treatment or patient sequestration), the PPV is indicative of an assay's ultimate utility and cost-effectiveness. When this metric is used, the real-time PCR test returned a PPV of only 65%. The LIBS classification tests reported here using the external validation criteria possessed PPVs of 95%, 60%, 92%, and 96% for the four genera shown in Table 2. These high PPVs are the result of very low rates of false positives, as can be seen in Table 2. By comparison, the gold-standard culture and count method has a PPV of 100% [19].

D. Use of Sequential Classification Tests to Increase Accuracy

The use of sequential chemometric models or algorithms, each based on the result of the previous test, to increase accuracy in the classification of LIBS spectra is beginning to gain acceptance. See, for example, Multari *et al.* [8]. To demonstrate the usefulness of this approach, we investigated the accuracy of identification of *E. coli* strains by using a two-step test. The first test classified an unknown spectrum using the five-genus library described earlier. We were interested in those spectra classified as *Escherichia* or *Enterobacter*, already noted as being highly similar. Bacterial spectra from those two genera were then further tested in a DFA model that only contained the *E. coli* spectra and the *E. cloacae* spectra. The result of this second classification was an increase in the *E. coli* identification relative to the five-genus library.

Specifically, 299 *E. coli* spectra were classified against the five-genus library. Again, an external validation was performed where entire sets of data were excluded from the analysis so that no other spectrum obtained from that data set was contained

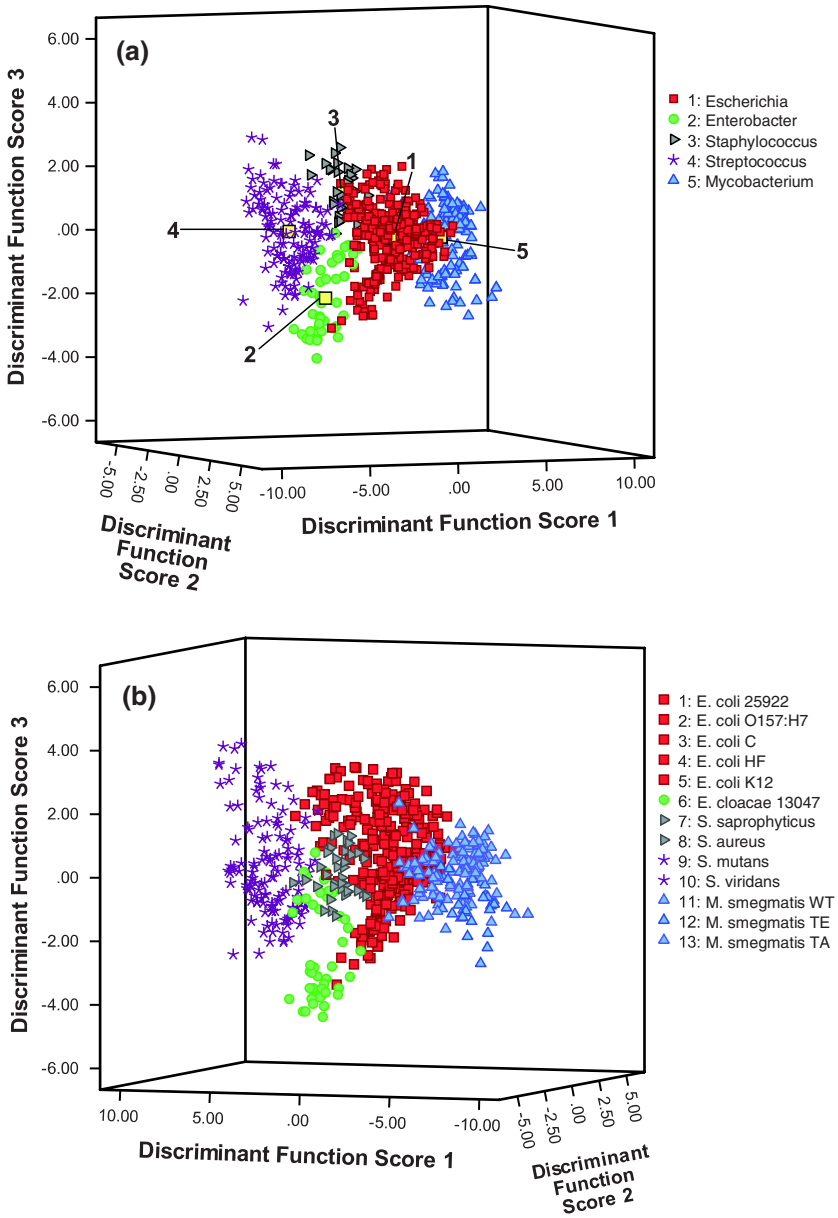

Fig. 3. (Color online) Spectral library. Plots showing the first three discriminant function scores from a DFA of 669 bacterial LIBS spectra collected in our laboratory over three years. (a) The spectra were classified into five distinct classes on the basis of bacterial genus. (b) The spectra were classified into 13 unique bacterial classes (see Table 1). The symbols of the 13 classes have been made uniform with (a) to show the similarities between the DFA when it is performed in these two different ways.

in the library. In this analysis, 269 of the *E. coli* spectra were correctly classified as *Escherichia* (89.97%), 21 spectra (7.02%) were incorrectly classified as *Enterobacter*, and the rest were incorrectly identified as another genus. When these 299 spectra were classified using a model containing only the *E. coli* and the *E. cloacae* spectra, 290 spectra (96.99%) were correctly classified as *E. coli* and only 9 (3.01%) were incorrectly classified at *E. cloacae*. This improvement of accuracy due to use of the more specialized model demonstrates that spectra that are identified as belonging to closely related organisms such as *Escherichia* and *Enterobacter*, or species within a given genus (i.e., the streptococci *viridans* or *mutans*), should be tested with a refined model specifically constructed to resolve the smaller subtle differences after a general genus-level classification has been made.

For completeness, the same sequential test was performed on the *E. cloacae* spectra. Because there was only one data set obtained from this organism, the external validation was not possible; therefore, samples were tested using the LOO cross validation. Using the five-genus model, 39 of 43 (90.70%) *E. cloacae* spectra were correctly classified as *E. cloacae*. The other four were identified as something other than *Escherichia* or *Enterobacter*. In the model containing only *E. coli* and *E. cloacae*, 40

Table 2. Truth Tables from a DFA Classification of 669 Bacterial LIBS Spectra in a Five-Genus DFA Model

External validation			Leave-one-out cross validation		
Escherichia	TRUE	FALSE	Escherichia	TRUE	FALSE
Positive	89.97%	4.28%	Positive	94.31%	0.61%
Negative	95.72%	10.03%	Negative	99.39%	5.69%
Staphylococcus	TRUE	FALSE	Staphylococcus	TRUE	FALSE
Positive	62.16%	2.55%	Positive	100.00%	0.51%
Negative	97.45%	37.84%	Negative	99.49%	0.00%
Streptococcus	TRUE	FALSE	Streptococcus	TRUE	FALSE
Positive	83.82%	2.04%	Positive	95.59%	0.00%
Negative	97.96%	16.18%	Negative	100.00%	4.41%
Mycobacterium	TRUE	FALSE	Mycobacterium	TRUE	FALSE
Positive	89.61%	1.27%	Positive	97.40%	0.00%
Negative	98.73%	10.39%	Negative	100.00%	2.60%

spectra (93.02%) were correctly classified as *E. cloacae* and three were classified as *E. coli*. This result, which shows little difference between the two models, demonstrates the fallibility of utilizing "LOO" cross-validation testing.

4. Summary

In this article we have reported the results of several experiments that demonstrate that a LIBS-based bacterial identification could be successfully applied to identify pathogens in a clinical specimen. The issue of mixed samples has once again been addressed with mixtures prepared from two closely related bacteria (*E. coli* and *E. cloacae*) using clinically relevant and realistic mixing fractions. This was intended to represent situations of clinical infections with the presence of additional background or contaminant bacteria.

To simulate the use of a LIBS-based test to diagnose urinary tract infections, cells of *S. epidermidis* were spiked into sterile urine samples and the bacteria were harvested from these samples with no other preparation or washing. Using a discriminant function analysis model that contained spectra from *S. epidermidis* cells harvested from water as well as two other staphylococci species, the urine-harvested samples were identified with 100% accuracy. No spectrum used to construct the model was obtained from cells obtained from a urine sample.

A 669-spectrum library composed of spectra from five bacterial genera and 13 distinct taxonomic groups was compiled and tested using external validation techniques where the model did not contain any spectra from samples acquired at the same time. Little difference was observed between the five-class model and the 13-class model. Truth tables constructed from the external validation of the five-genus model yielded sensitivities of approximately 85% and specificities above 95%. These external validation tests were compared to LOO cross-validation tests and, as expected, an artificially high accuracy was observed in the LOO tests.

Last, the use of sequential classification testing was investigated by identifying closely related *E. coli* and *E. cloacae* spectra using first the five-genus model spectral library and then a more targeted model that contained only *E. coli* and *E. cloacae* spectra. The improvement in accuracy obtained by using the more targeted model demonstrated the utility of using sequential testing to "filter" unknown spectra as they are tested, first through a "coarse" (perhaps genus-level) model and then through a more "fine" species-level (or similar) model to eventually obtain an accurate identification at the desired level. This simple addition of sequential testing can be easily implemented with little to no increase in complexity or time required.

The authors gratefully acknowledge the assistance of Dr. Hossein Salimnia and Dr. Robert Mitchell (DMC University Laboratories) for their helpful discussions, advice, and sample preparation. The authors also thank Ms. Laura Gunther for her assistance in preparing the mixture samples.

References

1. A. C. Samuels, F. C. DeLucia Jr., K. L. McNesby, and A. W. Miziolek, "Laser-induced breakdown spectroscopy of bacterial spores, molds, pollens, and protein: initial studies of discrimination potential," Appl. Opt. **42**, 6205–6209 (2003).
2. P. B. Dixon and D. W. Hahn, "Feasibility of detection and identification of individual bioaerosols using laser-induced breakdown spectroscopy," Anal. Chem. **77**, 631–638 (2005).
3. M. Baudelet, J. Yu, M. Bossu, J. Jovelet, J.-P. Wolf, T. Amodeo, E. Frejafon, and P. Laloi, "Discrimination of microbiological samples using femtosecond laser-induced breakdown spectroscopy," Appl. Phys. Lett. **89**, 163903 (2006).
4. M. Baudelet, L. Guyon, J. Yu, J.-P. Wolf, T. Amodeo, E. Frejafon, and P. Laloi, "Femtosecond time-resolved laser-induced breakdown spectroscopy for detection and identification of bacteria: a comparison to the nanosecond regime," J. Appl. Phys. **99**, 084701 (2006).
5. J. Diedrich, S. J. Rehse, and S. Palchaudhuri, "*Escherichia coli* identification and strain discrimination using nanosecond laser-induced breakdown spectroscopy," Appl. Phys. Lett. **90**, 163901 (2007).
6. J. Diedrich, S. J. Rehse, and S. Palchaudhuri, "Pathogenic *Escherichia coli* strain discrimination using laser-induced breakdown spectroscopy," J. Appl. Phys. **102**, 014702 (2007).
7. S. J. Rehse, Q. I. Mohaidat, and S. Palchaudhuri, "Towards the clinical application of laser-induced breakdown spectroscopy for rapid pathogen diagnosis: the effect of mixed cultures and sample dilution on bacterial identification," Appl. Opt. **49**, C27–C35 (2010).
8. R. Multari, D. A. Cremers, J. M. Dupre, and J. E. Gustafson, "The use of laser-induced breakdown spectroscopy for distinguishing between bacterial pathogen species and strains," Appl. Spectrosc. **64**, 750–759 (2010).

9. Q. Mohaidat, S. Palchaudhuri, and S. J. Rehse, "The effect of bacterial environmental and metabolic stresses on a LIBS-based identification of *Escherichia coli* and *Streptococcus viridans*," Appl. Spectrosc. **65**, 386–392 (2011).
10. D. Marcos-Martinez, J. A. Ayala, R. C. Izquierdo-Hornillos, F. J. de Villena, and J. O. Caceres, "Identification and discrimination of bacterial strains by laser induced breakdown spectroscopy and neural networks," Talanta **84**, 730–737 (2011).
11. C. A. Munson, F. C. DeLucia Jr., T. Piehler, K. L. McNesby, and A. W. Miziolek, "Investigation of statistics strategies for improving the discriminating power of laser-induced breakdown spectroscopy for chemical and biological warfare agent simulants," Spectrochim. Acta B **60**, 1217–1224 (2005).
12. M. Sabsabi, V. Detalle, M. A. Harith, W. Tawfik, and H. Imam, "Comparative study of two new commercial echelle spectrometers equipped with intensified CCD for analysis of laser-induced breakdown spectroscopy," Appl. Opt. **42**, 6094–6099 (2003).
13. B. Foxman, R. Barlow, H. D'Arcy, B. Gillespie, and J. D. Sobel, "Urinary tract infection: self-reported incidence and associated costs," Ann. Epidemiol. **10**, 509–515 (2000).
14. P. Whiting, M. Westwood, I. Watt, J. Cooper, and J. Kleijnen, "Rapid tests and urine sampling techniques for the diagnosis of urinary tract infection (UTI) in children under five years: a systematic review," BMC Pediatr. **5**, 1471–2431 (2005).
15. V. A. Fischetti, R. P. Novick, J. J. Ferretti, D. A. Portnoy, and J. I. Rood, *Gram-Positive Pathogens*, 2nd ed. (ASM, 2006).
16. N. De and M. Godlove, "Prevalence of *S. aureus* and *S. epidermidis* among patients with indwelling catheters and their antibiogram using some commonly used antibiotics," J. Am. Sci. **6**, 515–520 (2010).
17. D. E. Hall and J. A. Snitzer, III, "Staphylococcus epidermidis as a cause of urinary tract infections in children," J. Pediatr. **124**, 437–438 (1994).
18. M. Schaechter, N. C. Engleberg, B. I. Eisenstein, and G. Medoff, eds., *Mechanisms of Microbial Disease*, 3rd ed. (ASM, 1999).
19. L. O. Conterno, J. Shymanski, K. Ramotar, B. Toye, C. van Walraven, D. Coyle, and V. R. Roth, "Real-time polymerase chain reaction detection of methicillin-resistant *Staphylococcus aureus*: impact on nosocomial transmission and costs," Infect. Cont. Hosp. Ep. **28**, 1134–1141 (2007).