The use of Laser-Induced Breakdown Spectroscopy for the Differentiation of Pathogens and Viruses on Substrates

Rosalie A. Multari,[*] David A. Cremers, and Melissa L. Bostian

Applied Research Associates, Inc., 4300 San Mateo Blvd NE., Suite A-220, Albuquerque, NM 87110

[*]*Corresponding author: rmultari@ara.com*

In this work, the use of LIBS to differentiate live pathogens and killed viruses on substrates is investigated. Live pathogens *B. anthracis* Stern strain and *F. tularensis* live vaccine strain were interrogated as lawn and colonies on agar; dilutions on agar; and dilutions on glass slides and it was found possible to differentiate between all samples. UV killed *hantavirus* strains were studied as dilutions on slides and it was also found possible to differentiate between strains. To the best of our knowledge, this is the first study in which LIBS has been used to differentiate virus samples.

OCIS codes: 300.0300, 140.3440.

Introduction

Laser-induced breakdown spectroscopy (LIBS) is a spectroscopic analysis technique in which a laser pulse vaporizes ng to μg quantities of material and thermally excites the vaporized material in a short-lived plasma (~8000 K). Light emitted from atoms, ions, and simple molecules in the plasma is collected and analyzed. Traditionally, LIBS is an elemental analysis technique used to determine the composition of the target material via unique elemental fingerprint spectra. More

recently, chemometric [1] or other analysis techniques have been applied to LIBS spectra for both classification and identification of various materials.

In recent years, the use of Laser-Induced Breakdown Spectroscopy (LIBS) for the detection and differentiation of bacterial samples has become of interest. Gottfried et al. used LIBS to differentiate anthrax and ricin surrogates and nerve agent stimulants on substrates of metal and cardboard [2]. In other works, bacteria samples have been classified using atomic composition [3,4]. In previous work by us [5], pure lyophilized bacterial samples were differentiated using full spectrum analysis without specific correlation to atomic composition. More recently, Mohaidat et al. investigated the effect of adverse environmental and metabolic stress on LIBS identification of bacterial specimens and found it to be relatively robust [6].

In this work, the analysis method is to build chemometric models [1] to differentiate the samples and use these models in combination with a predictive flow for sample differentiation within the sample set. This method was used previously by us for the differentiation of live pathogen samples [5] and will herein be applied to the analysis of live pathogen samples in the form of lawn on agar ("lawn" is a uniform and uninterrupted layer of bacterial growth, in which individual colonies cannot be observed), colonies on agar, dilutions on agar, and dilutions on glass slides. This method will also be applied to dilutions of UV killed virus samples. Here, we use the term "dilution" to mean a pathogen colony sample or a viral sample mixed into a liquid such as water or Iodixanol.

It should be noted that the analysis presented does not demonstrate absolute detection (detection of the pure material) of the pathogen or virus under study as the LIBS spectra observed is the combined signal resulting from sampling the substrate, the pathogen or virus, the diluent (if present), and the atmosphere surrounding the sample. What is demonstrated is the

ability to differentially identify samples from within a predefined set of samples when they are presented and interrogated in the same manner. Differential analysis will be useful for LIBS instruments designed for specific applications such as manufacturing process control, detection of contaminants, and identifications for which the substrate material is the same or similar enough that a signal unique to the detection target can be extracted via mathematical analysis as opposed to applications in which absolute detection is desired.

Method

Apparatus and Samples

All samples used in this work were prepared by the University of New Mexico Health Sciences Center (UNMHSC) and data were collected on samples located inside a BSL-2 hood on-site at the center. The experimental arrangement used is shown in Figure 1 and described elsewhere [5]. The only difference here was the use of a dual channel spectrometer (Avantes) in place of a compact echelle spectrometer and ICCD. The LIBS data collection equipment was positioned outside the BSL-2 hood in which the samples were interrogated. For each sample, 100 LIBS spectra were collected with each individual spectrum being the average of 10 single shot spectra. Samples were moved manually during data collection to ensure data were collected from a fresh location on the sample for each shot.

The pathogen samples provided by UNMHSC consisted of live *B. anthracis* Sterne strain and *F. tularenis* live vaccine strain (*LVS*). Different agars were necessary for the growth of the pathogens. The *B. anthracis* samples were grown on blood agar (sheep blood) and the *LVS* samples were grown on cystine heart agar (rabbit blood). These agars are visibly different and care was taken so the person doing the data collection could not identify the sample being

analyzed. In addition, dilutions of the colony samples were prepared and spread on blood agar plates and glass slides by UNMHSC personnel to investigate LIBS differentiation capability when the substrate is the same (Sterne plated at 1 x 10^8 cfu/plate and *LVS* plated at 6 x 10^8 cfu/plate; cfu = colony-forming units). Pathogen data were also collected on an unknown blind sample set in the form of lawns on agar, dilutions on agar, and dilutions on slides. The *hantavirus* samples were diluted in Iodixanol and consisted of four different *hantavirus* strains: SNV, BAY, PUU, and RMV. For all dilutions, care was taken by UNMHSC personnel to ensure the same concentrations were applied for groups of samples to be analyzed together. "Blank" samples were also included in the analysis. For the live pathogens, the blanks were blood agar, cystine heart agar, and a clean glass slide. For the *hantavirus* strains, blanks were clean glass slides, water, and Iodixanol. See Table 1 for a summary of the samples included in this study. As the goal of this work was to demonstrate only differentiation of the different pathogens or viruses, no attempt was made here to ascertain the number of pathogens interrogated by each laser pulse. Relatively high concentrations were used to ensure the ability to differentiate. Determination of detection limits remains for a future study.

Spectral Analysis

The method of analysis used in this study is to build mathematical models to differentiate samples or groups of samples included in the analysis and then use these models in a predictive flow that relies on sequential screening [7]. The differentiation models are based on single variable partial least square regression combined with principal component analysis. This technique is commonly referred to as PLS and is especially useful when trying to predict a set of dependent variables from a very large set of independent variables. In PLS analysis it is assumed that all of the measured variance is useful variance to be explained. The latent variables are

estimated as exact linear combinations of the observed measures to create an exact definition of component scores. Through an iterative estimation technique, a general model is developed which encompasses canonical correlation, redundancy analysis, multiple regression, multivariate analysis of variance, and principal components. The iterative algorithm consists of a series of ordinary squares analyses. No distributional form is assumed for the measured variables. PLS1, used in this analysis, is a PLS method in which only one variable is modeled. Once a model has been generated for the sample classes, it can be used on test samples to produce a predictor value (in this case between 0 and 10) to be used to match the input tested sample to one of the sample classes. For this analysis, the dependent variable being the sample and the independent variables associated with the sample being the intensity measurements at each wavelength. Commercially available software (The Unscrambler™, Camo Software Inc.) was used for analysis and spectral normalization was applied to all spectra prior to building models. For the modeling of all samples except those on glass slides, the entire observed wavelength range was used as the variable set with the measured intensity treated as the values of the variables. Because the entire spectral range of 237–1015 nm was used, the modeling was done over 4096 variables (spectrometer channels) for each sample (dual channel spectrometer: 237-494 nm, 0.32 nm resolution; 495-1002 nm, 0.6 nm resolution). For the samples on glass slides, the spectral region of 577.28-598.67 nm was omitted from the analysis due to the presence of an intense Na emission line that added no value to the differential analysis and the modeling was done over 4019 variables.

For each group of samples to be differentiated, models were created using 50 of the 100 spectra collected for each sample and tested on the remaining 50 spectra reserved for the evaluation of model performance on input spectra. When a spectrum was input into a

differentiation model, the model produced a “prediction value” that indicated how well the input sample was matched to the two sample groups used to build the model. Good discrimination models were considered to be those that resulted in a wide separation of the prediction values for the two groups. A value in the prediction value gap was then chosen for which input samples with higher prediction values would be considered matched to the sample being discriminated. Samples with prediction values lower than this chosen prediction value would be considered matched to the samples not being discriminated. Figure 2 illustrates this process. As shown in the figure, the best models are those for which there is a wide separation in the prediction values obtained from the model for the two sample groups being differentiated when test spectra are input. To improve the observed separation, prediction values obtained for individual spectra were averaged as needed. For the agar samples, it was necessary to average predictions for 50 spectra to get a separation that could be used for sample screening through a predictive flow. For the samples on glass slides, it was only necessary to average prediction values from 10 spectra. This implies that there is less variability in the LIBS spectra when the sample is interrogated on a glass slide compared to interrogation on agar.

Results and Discussion

Pathogens

Using the method of analysis described above, data collected for the pathogen samples were analyzed in three separate differential groups: lawn on agar, dilution on agar, and dilutions on glass slides. Uncontaminated or “blank” sample differentiation models were also created to identify uncontaminated agar samples.

As can be seen for the lawn sample prediction results in Figure 3, good differentiation was achieved between the sample groups and it was possible to build single models capable of differentiating both types of the blank agar (cystine heart and blood agar), *B. anthracis*, and *LVS* from all other samples with as few as 10 prediction values averaged for screening. By inputting the unknown sample spectra into these models, it is clear that none of the "unknown" samples are uncontaminated agar and that L4 is a lawn sample of *B. anthracis* and L2 and L3 are lawn samples of *LVS*. It should be noted that the substrate is not the same for *B. anthracis* and *LVS* and this difference may be contributing to the robustness of the differentiation models created. All lawn "unknowns" were correctly identified.

For the dilution on agar samples, it was only necessary to build a single model to differentiate *B. anthracis* from *LVS* as the dilution samples were previously screened to see if any were uncontaminated agar. The results of the model when test spectra and unknown spectra were input are shown in Figure 4. The model created for the differentiation of dilutions on agar produced prediction values with greater scatter than was observed for the lawn differentiation samples and all 50 prediction values obtained when the model was run on the test spectra were averaged to obtain the separation needed for sample identification. Even with averaging a larger number of prediction values, the separation between samples is not as good as the separation observed in the lawn sample analysis. In addition, the prediction values for all blind samples are lower than expected from the calibration sample data. *B. anthracis* samples should have prediction values of approximately 7 and *LVS* samples should have values of about 5.2. Knowledge that the blinds contained both *B. anthracis* and *LVS* indicated the prediction value distributions had shifted. Because the prediction values have shifted, it is difficult to unambiguously identify the blind samples. Based on our experience, shifts in prediction value

distributions are typically observed in cases for which an experimental parameter has changed (ex. sample concentration, coupling of laser energy into sample, etc.) and the variation that was introduced into the spectral shape by the change was not captured in the modeling. These results indicate that (1) more robust calibration is needed to identify the blind dilution samples (e.g. greater number of spectra collected) and (2) that substrate media and sampling methods must be carefully chosen to stabilize the prediction values and to separate the prediction values for the different samples to as great an extent as possible. For the purposes of this study, this experience was applied and the prediction value was adjusted to 4 to correctly report the identity of the blinds to UNMHSC. However, for the purposes of creating LIBS instrumentation, this differentiation model could not be used for sample differentiation and would have to be re-worked using additional data until all variability has been captured and the separation between prediction values is increased.

For the dilutions on slides, it was possible to build single differential models to differentiate the sample of interest from the other samples. As can be seen from Figure 5, good separation was again achieved between the sample groups by averaging only 10 prediction values obtained from 10 of the test spectra and the shift observed in the predictive distribution for these same samples when agar was used as the substrate is no longer present. From the predictive data, it is clear that none of the dilutions on slides are uncontaminated, and that samples A and C are *B. anthracis* dilutions on slides while samples B and D are *LVS* dilutions on slides. All dilution unknowns on slides were correctly identified. Further, it appears that slides make a better substrate for pathogen dilution analysis than does blood agar for the pathogens studied. In summary, all unknown pathogens included in this study were correctly identified.

Viruses

The method of analysis described above was next applied to LIBS spectra collected on four different strains of *hantavirus* in Iodixanol dilution applied to slides and allowed to dry, as well as a clean slide, and a slide with clean water on it (dried). It should be noted that Iodixanol is a patented isosmolar contrast medium available for intravascular use formulated with sodium and calcium in a ratio equivalent to blood and is often used to purify viruses and investigate their assembly [8, 9]. The UV killed hantavirus strains were: PUU, BAY, SNV, and RMV. As this was the first known attempt to determine if LIBS spectra could be used to differentiate virus samples, no unknowns were included in the study.

Various models were attempted and it was not possible to differentiate these samples using single models for differentiation so a predictive flow approach with sequential differentiation was implemented [7]. In the predictive flow approach, the first and easiest sample to separate from the group was the blank slide sample. After the blank slide spectra were removed from the analysis, the next easiest samples to differentiate were the Iodixanol only sample followed by the PUU dilution sample (see Figure 6). Once these three samples were removed from the differential grouping analysis, it was then possible to build single models for the differentiation of the remaining three *hantavirus* strains (BAY, SNV, and RMV).

The nature of the spectral differences between the four strains of *hantavirus* that permitted differentiation using LIBS was not determined in this study. It is known, however, that strains of other viruses (e.g. Tobacco Mosaic Virus) [10] contain different percentages of certain amino acids which may be the source of differentiation here. Examination of the regression model loadings from our virus study shows that the greatest variances occur for the elements: C (247 nm), Mg (279.6 and 280.3 nm), Ca (393.3, 396.8, and 422.6 nm), H (656.3 nm), N (742.3, 744.2,

746.8, 822.3, and 868.0 nm), and O (777.4 and 844.6 nm) which, with the exception of Ca and Mg which are probably contaminants, are the main elements in amino acids. Further work is required, however, to fully understand the biological differences between the samples resulting in sufficient spectral variation to make differentiation between sample groups possible.

Conclusions

From this study it can be concluded that differential analysis of integrated LIBS spectra (substrate, target species, and atmosphere) of the pathogens and viruses used here is possible. This suggests the future potential use of LIBS as a bio-diagnostic tool for applications in which key analysis parameters of pathogen, substrate, and measurement conditions are well defined. Much work remains, however, to bridge the gap between laboratory findings such as those presented here and a LIBS instrument ready for use in the "real" world. Further work is also required to fully understand the biological differences between the samples resulting in sufficient spectral variation to make differentiation between sample groups possible. In addition, it has been shown that for the pathogen and viruses used here that glass slide substrates produced better differential modeling results than sampling on agar substrates. This finding indicates that the choice of substrate can strongly impact the success of differentiation and that a careful study of the best sampling method to produce optimal differential modeling must be a critical part of the creation of a LIBS based diagnostic tool.

Acknowledgements

The authors wish to acknowledge and thank the University of New Mexico Health Sciences Center for preparing and providing the samples used for this study as well as the use of their bio

safe facilities. This work was support by Applied Research Associates, Inc. under project IRAD R00074.

REFERENCES

1. K.H. Esbensen, *Multivariate Data Analysis-In Practice, 5th ed*. (Camo, Oslo, Norway, 1994).
2. J.L. Gottfried, F.C. De Lucia, Jr., C.A. Munson, and A.W. Miziolek, "Standoff Detection of Chemical and Biological Threats Using LIBS", Appl. Spectrosc., **62**, No. 4, 353-363 (2008).
3. J. Diedrich, S.J. Rehse, and S. Palchaudhuri, "Pathogenic Escherichia coli strain discrimination using laser-induced breakdown spectroscopy," J. Appl. Phys. **102**, 014702 (2007).
4. 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).
5. R.A. Multari, D.A. Cremers, J.M. Dupre, J.E. Gustafson, "The Use of Laser-Induced Breakdown Spectroscopy for Distinguishing Between Bacterial Pathogen Species and Strains," Appl. Spectrosc. **64**, 750-759 (2010).
6. 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).
7. United States of America Patent Application 12/981,626.
8. E. Gias, S.U. Nielsen, S.U., L.A.F. Morgan,. and G.L. Toms, " Purification of human respiratory syncytial virus by ultracentrifugation in iodixanol density gradient," J. Virol. Meth., **147**, 328-332 (2008).

9. J.B. Prescott, P.R. Hall, V.S. Bondu-Hawkins, C. Ye, C. and B. Hjelle, “Early innate immune responses to Sin Nombre Hantavirus occur independently of IFN regulatory factor 3, characterized pattern recognition receptors andviral entry,” J. Immunol., **179**, 1796-1802 (2007).
10. C.A. Knight, “The Nature of Some of the Chemical Differences Among Strains of Tobacco Mosaic Virus,” Presented in Part before the Division of Biological Chemistry at the American Chemical Society at Atlantic City, April 14-18, 1947. [Available at : http://www.jbc.org/content/171/1/297.full.pdf; last accessed 11-November 2011]

FIGURE CAPTIONS

Figure 1. Experimental set-up used to collect LIBS spectra of pathogens and viruses. The samples were located inside a BSL-2 hood. LIBS emission was collected along the path of the laser light to remove parallax.

Figure 2. Illustration of the process used to create the differential models needed for sample screening. The best models are those in which there is a wide separation in the prediction values obtained for the two sample groups being differentiated as is shown below for the differentiation of *B. anthracis* from *LVS* in dilution on slides. Example spectra used to create a differentiation model are shown on the left. On the top right the two dimensional score space plot for the resulting model is shown. In the lower right, the prediction value results (average 20 spectra) are shown obtained for spectra reserved for testing the model.

Figure 3. Prediction results for differentiation models built for the lawn samples when test spectra are input (left) and when unknown or "blind" spectra are input (right). The differential models in the first two rows were designed to differentiate uncontaminated agar samples from all other samples (including dilutions on agar) whereas the models in the rows below were designed to differentiate the *B. anthracis* sample and the *LVS* sample from all other samples.

Figure 4. Prediction results for differentiation models built for the dilution samples when test spectra are input (left) and when unknown or "blind" spectra are input (right). Note the predictive distribution was shifted for the unknown samples. This indicates sample variability was not successfully captured in the modeling. See text for a detailed discussion of the shifted distribution.

Figure 5. Prediction results for differentiation models built for the dilution on slide samples when test spectra are input (left) and when unknown or "blind" spectra are input (right). The

differential model in the top row was designed to differentiate uncontaminated slide samples from all other samples whereas the models in the rows below were designed to differentiate the *B. anthracis* dilution on slide sample from all other slide samples and the *LVS* dilution on slide sample from all other slide samples respectively.

Figure 6. Prediction results for differentiation models built for the *hantavirus* dilution on slide samples when test spectra are input. In the top row, prediction results on test spectra for models designed to differentiate the blank slide with water and the slide with Iodixanol are presented. The remaining rows show test spectra results for the *hantavirus* dilution models. Once the blank slide, the Iodixanol, and the PUU samples were removed from the analysis, it was possible to build single models to differentiate the remaining samples.

FIGURES

Figure 1. Experimental set-up used to collect LIBS spectra of pathogens and viruses. The samples were located inside a BSL-2 hood. LIBS emission was collected along the path of the laser light to remove parallax.

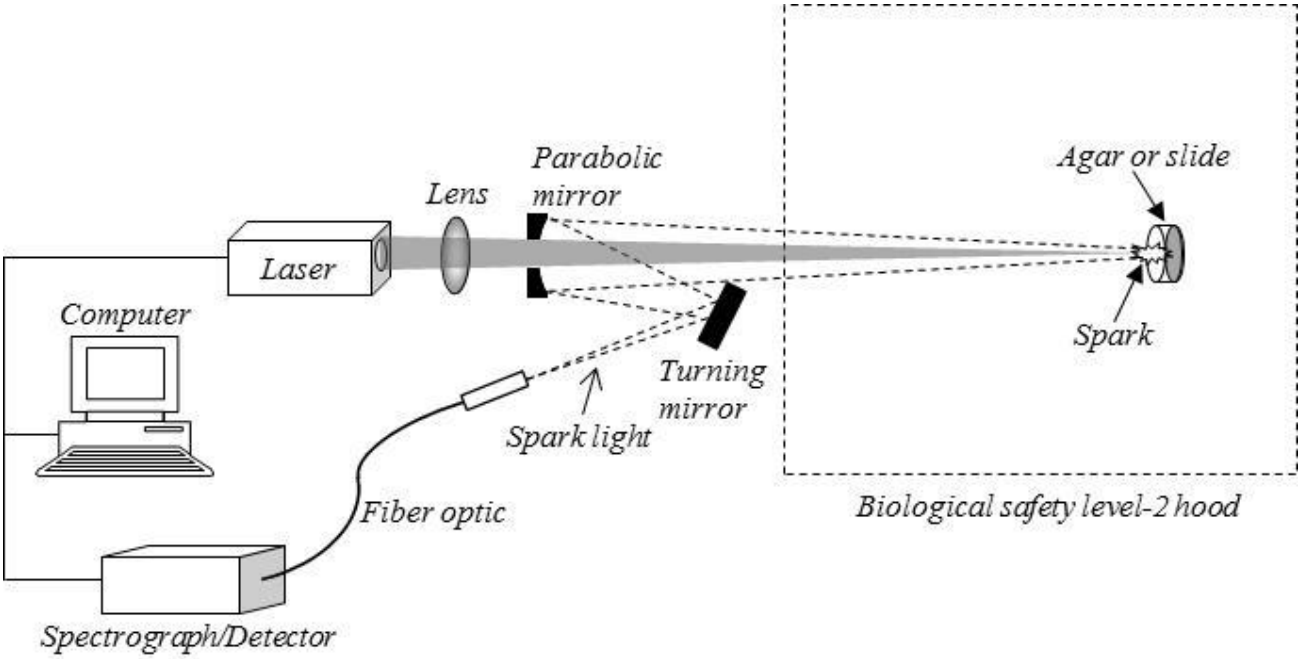


Figure 2. Illustration of the process used to create the differential models needed for sample screening. The best models are those in which there is a wide separation in the prediction values obtained for the two sample groups being differentiated as is shown below for the differentiation of *B. anthracis* from *LVS* in dilution on glass slides. Example spectra used to create a differentiation model are shown on the left. On the top right the two dimensional score space plot for the resulting model is shown. In the lower right, the prediction value results (average 20 spectra) are shown obtained for spectra reserved for testing the model.

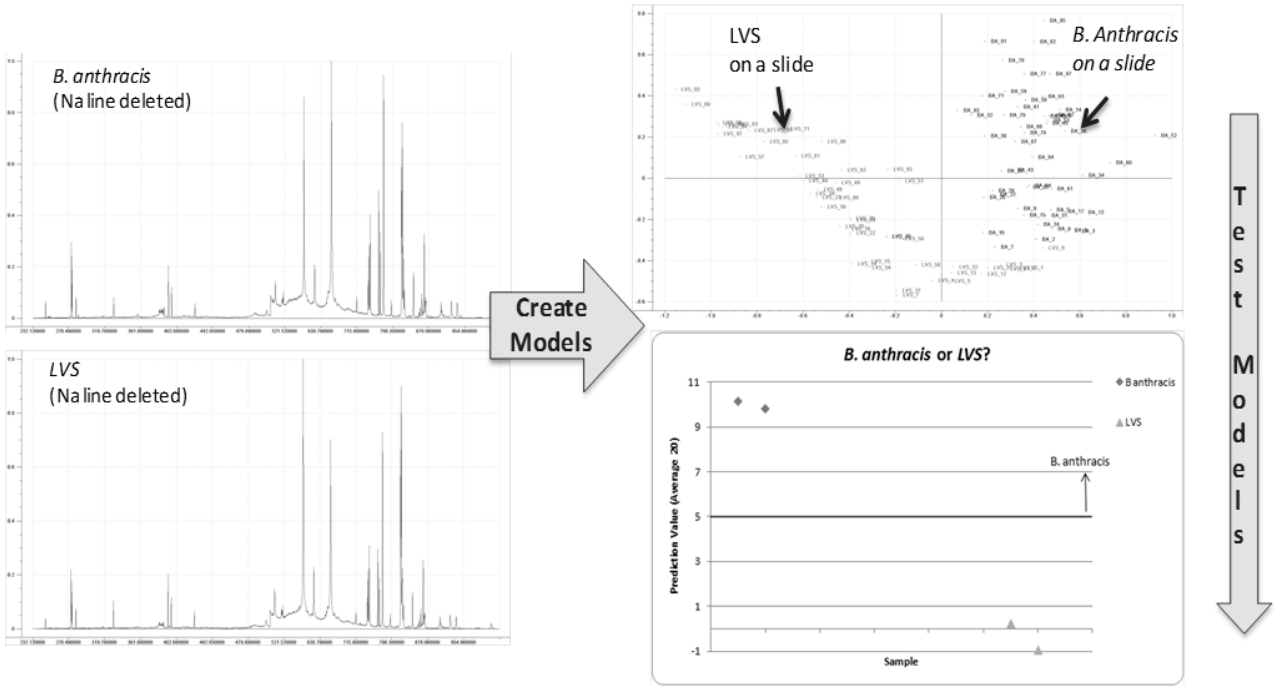


Figure 3. Prediction results for differentiation models built for the lawn samples when test spectra are input (left) and when unknown or "blind" spectra are input (right). The differential models in the first two rows were designed to differentiate uncontaminated agar samples from all other samples (including dilutions on agar) whereas the models in the rows below were designed to differentiate the *B. anthracis* sample and the *LVS* sample from all other samples.

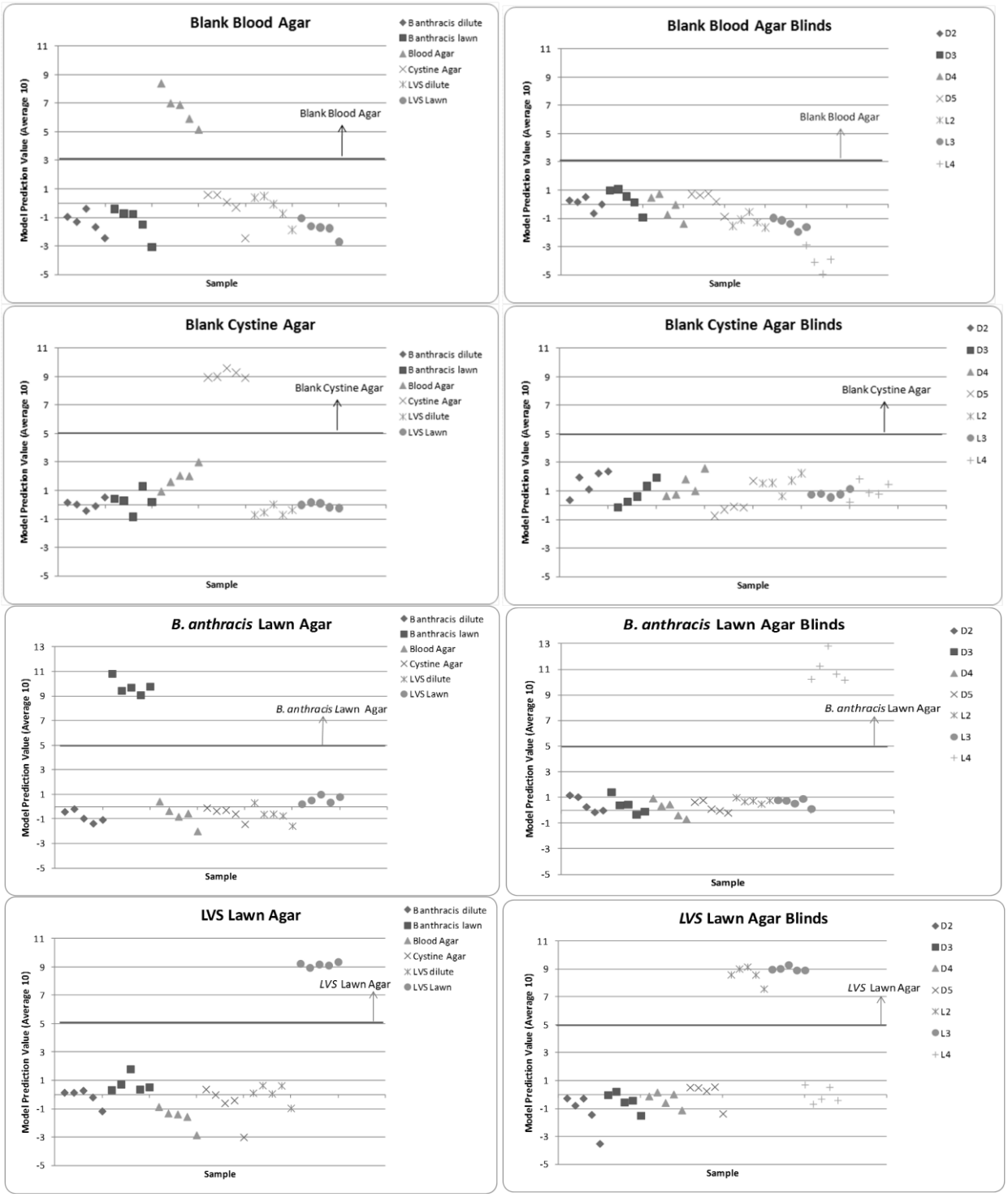


Figure 4. Prediction results for differentiation models built for the dilution samples when test spectra are input (left) and when unknown or "blind" spectra are input (right). Note the predictive distribution was shifted for the unknown samples. This indicates sample variability was not successfully captured in the modeling. See text for a detailed discussion of the shifted distribution.

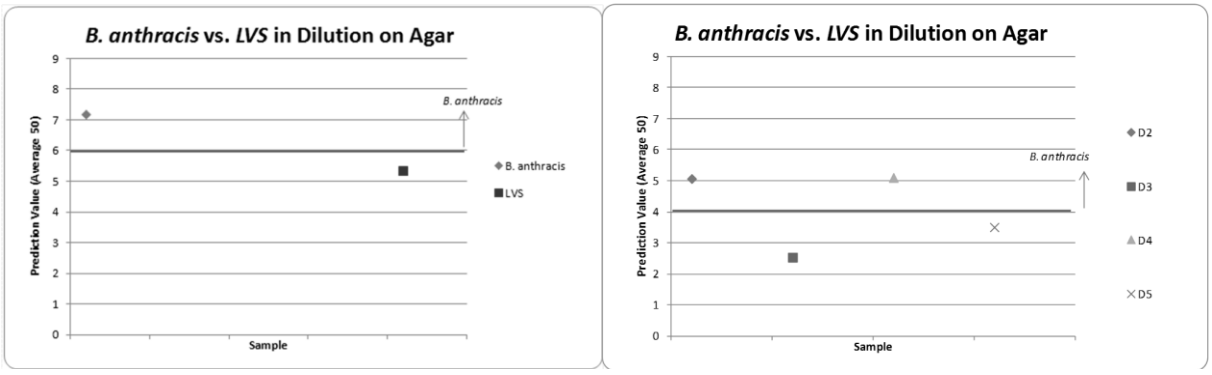


Figure 5. Prediction results for differentiation models built for the dilution on slide samples when test spectra are input (left) and when unknown or "blind" spectra are input (right). The differential model in the top row was designed to differentiate uncontaminated slide samples from all other samples whereas the models in the rows below were designed to differentiate the *B. anthracis* dilution on slide sample from all other slide samples and the *LVS* dilution on slide sample from all other slide samples respectively.

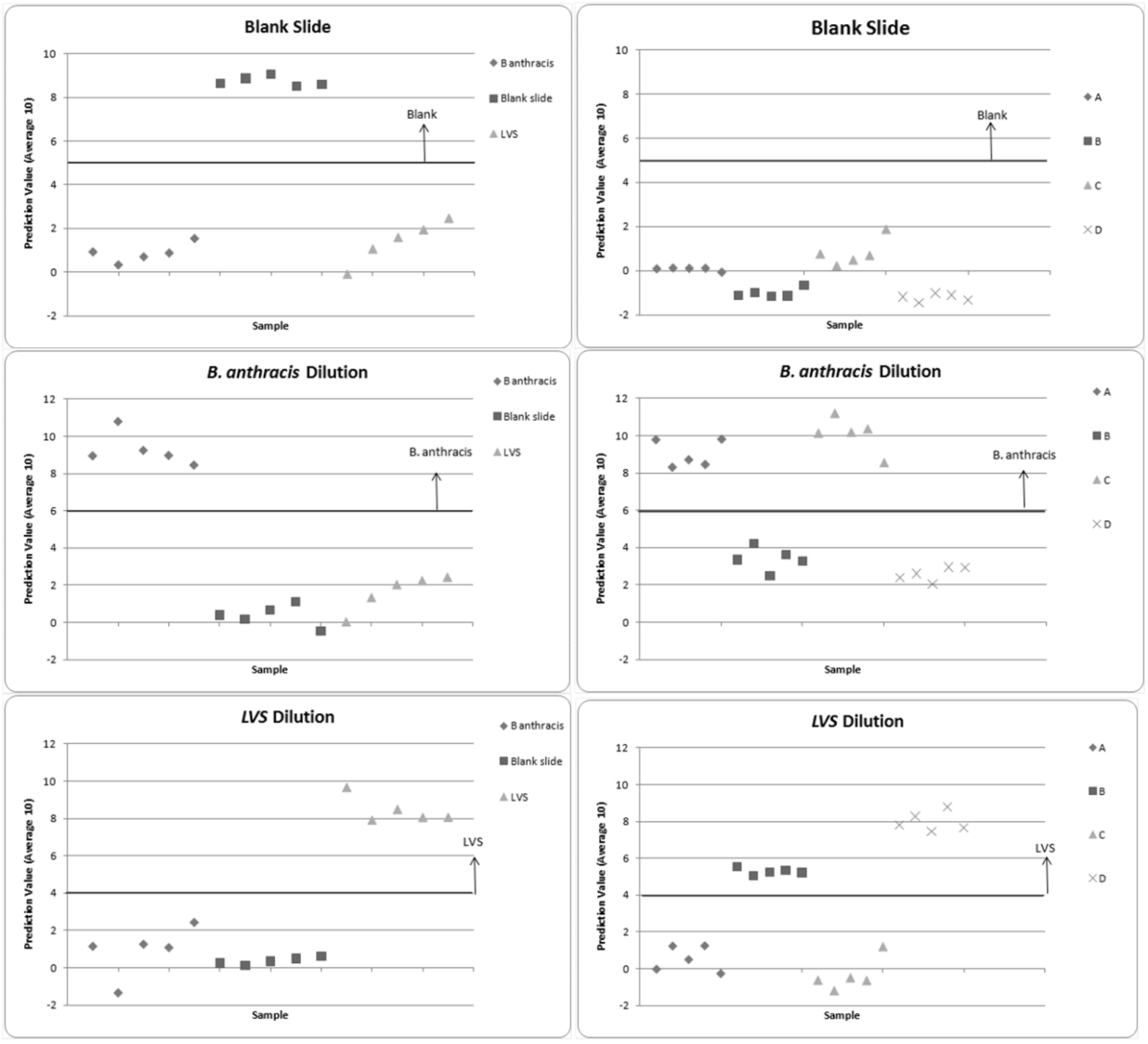
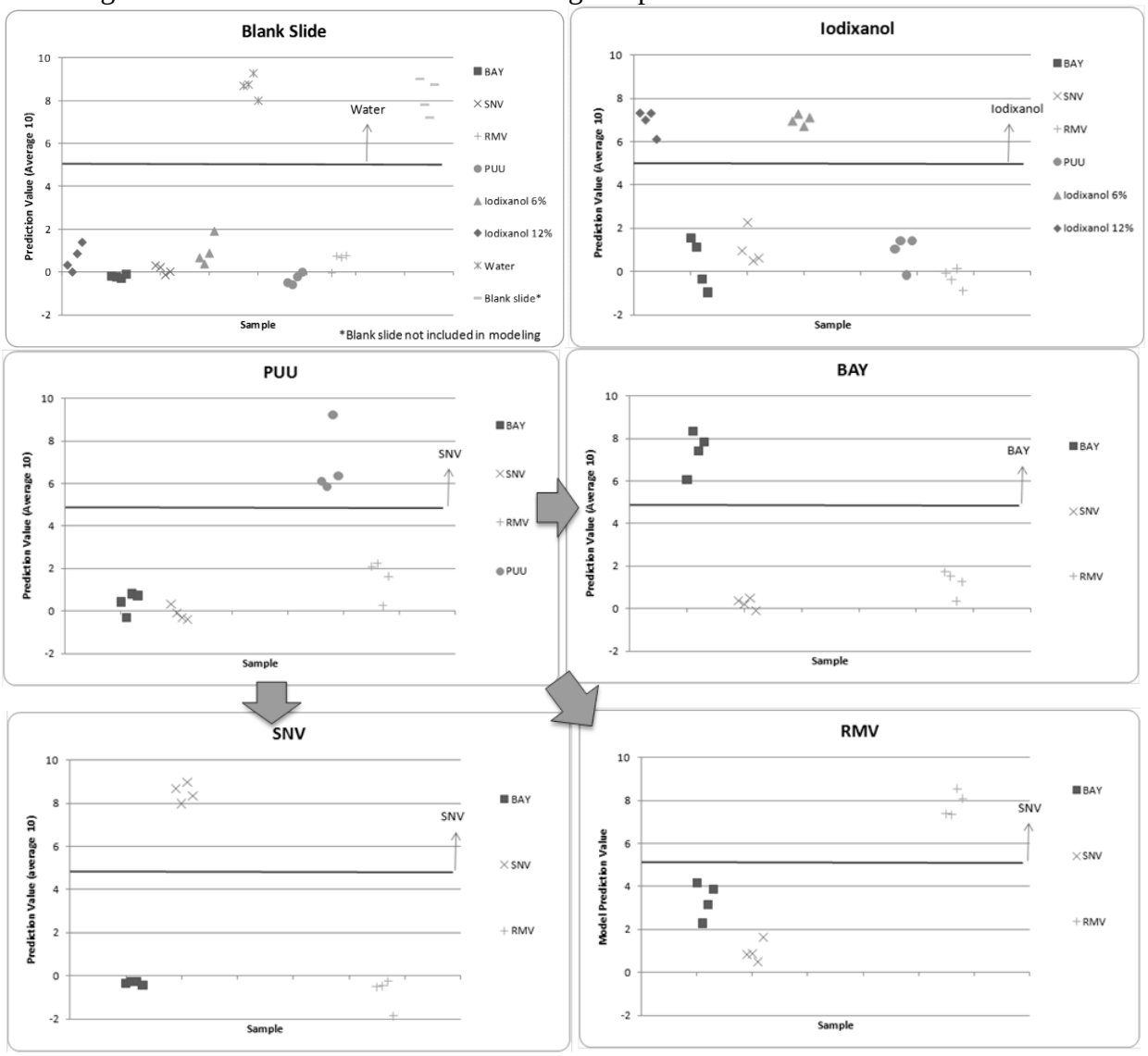


Figure 6. Prediction results for differentiation models built for the *hantavirus* dilution on slide samples when test spectra are input. In the top row, prediction results on test spectra for models designed to differentiate the blank slide with water and the slide with Iodixanol are presented. The remaining rows show test spectra results for the *hantavirus* dilution models. Once the blank slide, the Iodixanol, and the PUU samples were removed from the analysis, it was possible to build single models to differentiate the remaining samples.



TABLE

Table 1. Summary of samples prepared by the UNMHSC for this study

Sample	Substrate	Form	Metabolic State
B. anthracis Sterne strain	blood agar	lawn	live
B. anthracis Sterne strain	blood agar	dilution	live
B. anthracis Sterne strain	glass slide	dilution	live
F. Tularensis vaccine strain	cystine heart agar	lawn	live
F. Tularensis vaccine strain	blood agar	dilution	live
F. Tularensis vaccine strain	glass slide	lawn	live
Blank	blood agar	N/A	N/A
Blank	cystine heart agar	N/A	N/A
Blank	slide	N/A	N/A
L2, L3, L4	unknown agar	unknown lawn	live
D2, D3, D4, D5	blood agar	unknown dilution	live
A, B, C, D	glass slide	unknown dilution	live
Hantavirus (SNV)	glass slide	dilution	UV killed
Hantavirus (BAY)	glass slide	dilution	UV killed
Hantavirus (PUU)	glass slide	dilution	UV killed
Hantavirus (RMV)	glass slide	dilution	UV killed
Water	glass slide	N/A	N/A
Iodixanol	glass slide	N/A	N/A