Spectrochimica Acta Part B 87 (2013) 161–167

Contents lists available at ScienceDirect

Spectrochimica Acta Part B

journal homepage: www.elsevier.com/locate/sab

A comparison of multivariate analysis techniques and variable selection strategies in a laser-induced breakdown spectroscopy bacterial classification

Russell A. Putnam [a,1], Qassem I. Mohaidat [b], Andrew Daabous [a,1], Steven J. Rehse [a,*]

[a] Department of Physics, University of Windsor, Windsor, Ontario N9B 3P4, Canada
[b] Department of Physics, Yarmouk University, Irbid 21163, Jordan

ARTICLE INFO

Article history:
Received 15 November 2012
Accepted 12 May 2013
Available online 20 May 2013

Keywords:
Laser-induced breakdown spectroscopy
Bacterium
Discriminant function analysis
Partial least squares discriminant analysis
Multivariate analysis

ABSTRACT

Laser-induced breakdown spectroscopy has been used to obtain spectral fingerprints from live bacterial specimens from thirteen distinct taxonomic bacterial classes representative of five bacterial genera. By taking sums, ratios, and complex ratios of measured atomic emission line intensities three unique sets of independent variables (models) were constructed to determine which choice of independent variables provided optimal genus-level classification of unknown specimens utilizing a discriminant function analysis. A model composed of 80 independent variables constructed from simple and complex ratios of the measured emission line intensities was found to provide the greatest sensitivity and specificity. This model was then used in a partial least squares discriminant analysis to compare the performance of this multivariate technique with a discriminant function analysis. The partial least squares discriminant analysis possessed a higher true positive rate, possessed a higher false positive rate, and was more effective at distinguishing between highly similar spectra from closely related bacterial genera. This suggests it may be the preferred multivariate technique in future species-level or strain-level classifications.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Since the initial demonstrations of bacterial identification with laser-induced breakdown spectroscopy (LIBS) in 2003, significant progress has been made in the use of multivariate chemometric analyses to classify unknown bacterial LIBS spectra [1–4]. Over the last five years we and others have demonstrated a sensitive and specific identification of live bacterial biospecimens utilizing a discriminant function analysis (DFA) to classify LIBS spectra [5–8]. The intensities of strong specific elemental atomic emission lines normalized by the total observed spectral power have been utilized as independent variables in this multivariate analysis [9]. The selection of specific spectral lines to serve as independent variables in the multivariate analysis is known as variable down-selection [10]. However it is not yet known whether the use of down-selected variables or the entire LIBS spectrum provides optimal discrimination and classification of unknown LIBS spectra, and this is an ongoing area of investigation [11,12]. It is also not known which multivariate analysis technique, if any, provides superior classification given a choice of independent variables, and multiple chemometric algorithms are still widely utilized for bacterial identification including principal component analysis (PCA), linear discriminant analysis (LDA), partial least squares discriminant analysis (PLS-DA), neural network (NN) analysis, partial least squares (PLS) regression, and support vector machine classification (SVM) [13–18].

To investigate these various strategies, we have compared the use of three different down-selected variable models consisting of emission intensities, the sum of observed intensities from the elements P, Ca, Mg, Na, and C, and complex ratios of those intensities in identical external validation tests. Variables were down-selected from bacterial LIBS spectra obtained from five different genera and 13 distinct taxonomic classes of species and strains [8]. Model performance was quantified by calculating truth tables (and the resulting sensitivity and specificity) from the external validation tests. Lastly, the down selected variable model which provided the most accurate classification was tested in a PLS-DA multivariate analysis to provide a direct comparison with the performance of the DFA.

2. Experimental

2.1. Experimental setup

The LIBS apparatus used to obtain the bacterial spectra, as well as our bacterial sample preparation and mounting protocols, have been described at length elsewhere [5,19]. Briefly, 1064 nm infrared laser pulses 10 ns in duration were used to ablate the bacterial specimens mounted on a 0.7% nutrient-free agar substrate in an argon environment. LIBS emission was collected 2 μs after the ablation pulse and dispersed in an Échelle spectrograph, and the spectra were recorded

* Corresponding author. Tel.: +1 519 253 3000; fax: +1 519 973 7075.
E-mail addresses: putnamr@uwindsor.ca (R.A. Putnam), q.muhaidat@yu.edu.jo (Q.I. Mohaidat), daabousa@uwindsor.ca (A. Daabous), rehse@uwindsor.ca (S.J. Rehse).
[1] Fax: +1 519 973 7075.

0584-8547/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sab.2013.05.014

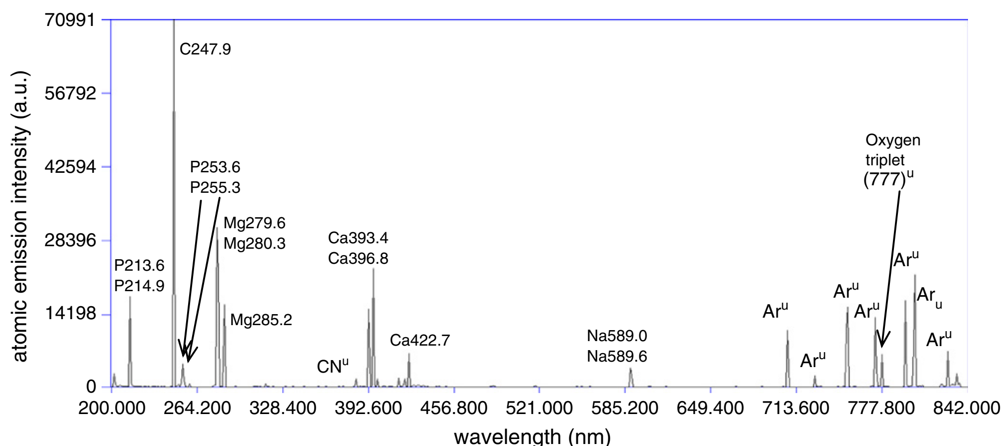

Fig. 1. A representative LIBS spectrum of a bacterial target ablated in an argon environment at atmospheric pressure. The atomic emission lines used in the bacterial discrimination indicated by an "*" in Table 3 are indicated in this spectrum. Emission features that were seen but were unused in the discrimination are indicated with a superscript "u".

by an intensified charge-coupled device (ESA3000, LLA Instruments, GmbH). Pulse energies were approximately 10 mJ/pulse and each spectrum was averaged from spectra acquired at five sampling locations, each approximately 100 μm in diameter. Approximately 7500 bacterial cells in total were ablated for each spectrum [5]. A representative LIBS spectrum of a bacterial target ablated on an agar substrate in an argon atmosphere is shown in Fig. 1. This spectrum is the averaged accumulation of five separate sampling locations. Five spectra were acquired at each sampling location, thus twenty-five laser pulses were used to obtain this spectrum.

The bacteria were chosen to represent a fairly wide taxonomic range. Spectra were acquired from representative Gram-negative phenotypes (*Escherichia coli* and *Enterobacter cloacae*), Gram-positive phenotypes (two species of *Staphylococci* and two species of *Streptococci*), and the atypical acid-fast *Mycobacterium* phenotype (three strains of *Mycobacterium smegmatis*). In total, LIBS spectra from 13 unique bacterial strains were obtained in 32 completely distinct experiments (e.g. cultured in different media, grown on different days over the course of 18 months, and exposed to different environmental stresses) [8]. This is shown in Table 1.

The five representative bacterial genera that were tested are listed in the first column of Table 1 and the thirteen bacterial taxonomic groups tested (e.g. *E. coli* strain C, *E. coli* strain HF4714, *Staphylococcus aureus*, *Staphylococcus saprophyticus*) are listed in column two. The 32 distinct experiments that were performed yielded the 32 data sets shown in column three of Table 1. Each distinct experiment was

Table 1
Identities of the 32 data sets used to construct a spectral library 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

Download English Version:

https://daneshyari.com/en/article/1239925

Download Persian Version:

https://daneshyari.com/article/1239925

Daneshyari.com