EDGEWOOD CHEMICAL BIOLOGICAL CENTER
U.S. ARMY RESEARCH, DEVELOPMENT AND ENGINEERING COMMAND
Aberdeen Proving Ground, MD 21010-5424

ECBC-TR-1460

MASS SPECTROMETRY PROTEOMICS METHOD AS A RAPID SCREENING TOOL FOR BACTERIAL CONTAMINATION OF FOOD

Rabih Jabbour
RESEARCH AND TECHNOLOGY DIRECTORATE

Havas, Karyn A.
U.S. DEPARTMENT OF AGRICULTURE
ANIMAL PLANT HEALTH INSPECTION SERVICES
FOREIGN ANIMAL DISEASE DIAGNOSTIC LAB
Riverdale, MD 20737-1230

Mary M. Wade
RESEARCH AND TECHNOLOGY DIRECTORATE

Samir V. Deshpande
SCIENCE AND TECHNOLOGY CORPORATION
Edgewood, MD 21040-2734

Patrick McCubbin
OPTIMETRICS INCORPORATED
Abingdon, MD 21009-1283

Candelaria C. Daniels
Bernardo Delgado
U.S. ARMY PUBLIC HEALTH COMMAND REGION-SOUTH
DOD FOOD ANALYSIS AND DIAGNOSTIC LABORATORY
Ft. Sam Houston, TX 78234-7583

June 2017

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
XX-06-2017	Final	Mar 2010 – Dec 2011

4. TITLE AND SUBTITLE
Mass Spectrometry Proteomics Method as a Rapid Screening Tool for Bacterial Contamination of Food

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jabbour, Rabih (ECBC); Havas, Karyn A. (USDA); Wade, Mary M. (ECBC); Deshpande, Samir V. (STC); McCubbin, Patrick (Optimetrics); Daniels, Candelaria C.; and Delgado, Bernardo (USAPHC-South)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Director, ECBC, ATTN: RDCB-DRI-D, APG, MD 21010-5424
USDA APHIS, 4700 River Road, Riverdale, MD 20737-1230
Science and Technology Corporation, 500 Edgewood Road, Suite 205, Edgewood, MD 21040-2734
Optimetrics, Inc. (A DCS Company), 100 Walter Ward Boulevard, Suite 100, Abingdon, MD 21009-1283
USAPHC-South FADL, 2899 Schofield Road, Suite 2630, Fort Sam Houston, TX 78234-7583

8. PERFORMING ORGANIZATION REPORT NUMBER
ECBC-TR-1460

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Command, 2748 Worth Road, Fort Sam Houston, TX 78234-7583

10. SPONSOR/MONITOR'S ACRONYM(S)
MEDCOM

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT:

Every year in the United States there are over 48 million cases of foodborne illness. Traditional microbiological techniques require multiple enrichments using selective media for pathogen detection. Accurate identification of the offending pathogen is necessary to provide the most appropriate outbreak response and patient care. The mass spectrometry proteomics method (MSPM) does not require enrichment and is not affected by pathogens. The ability to use the MSPM to correctly classify whether or not food samples were contaminated with *Salmonella enterica* serotype Newport in this blinded pilot study resulted in a high level of sensitivity and specificity (>99 and 98.6%, respectively). The study involved mashed potato samples spiked at a concentration of 10^6 cfu/mL. These initial studies are encouraging and require further evaluation in more complex food matrices and at various pathogen concentrations to validate MSPM as a useful foodborne pathogen diagnostic tool.

15. SUBJECT TERMS
Pathogen Mass spectrometry proteomics method (MSPM) Foodborne Food matrix

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Renu B. Rastogi
U	U	U	UU	26	**19b. TELEPHONE NUMBER** *(include area code)* (410) 436-7545

Blank

PREFACE

The work described in this report was started in March 2010 and completed in December 2011.

The use of either trade or manufacturers' names in this report does not constitute an official endorsement of any commercial products. This report may not be cited for purposes of advertisement.

This report has been approved for public release.

Acknowledgment

Funding for this project was provided by the U.S. Army Medical Command, Ft. Sam Houston, TX.

Blank

CONTENTS

FIGURES

TABLES

MASS SPECTROMETRY PROTEOMICS METHOD AS A RAPID SCREENING TOOL FOR BACTERIAL CONTAMINATION OF FOOD

1. INTRODUCTION

Food defense is a growing field and is necessary to protect populations from intentional adulteration of foodstuffs. Nefarious individuals can and have intentionally contaminated food sources using biological warfare agents and other pathogens. Examples of this include intentional contamination of salad bars and ground beef, sometimes committed by restaurant workers (Centers for Disease Control and Prevention, 2003; Kolavic et al., 1997; Török et al., 1997). In 2007, the U.S. Food and Drug Administration acknowledged the need for food defense and issued a *Food Protection Plan* to mitigate intentional food contamination (Food and Drug Administration, 2007). This intentional threat exists on top of the already high burden of diseases associated with accidental contaminations due to naturally occurring foodborne pathogens. In the United States, it is estimated that more than 9 million foodborne illnesses from identified pathogens are acquired each year from aquatic and land animals and plants (Painter et al., 2013; Scallan et al., 2011a, 2011b). Additional illnesses from foodborne disease caused by unspecified agents have been estimated at 38.4 million, which totals approximately 48 million cases of foodborne illness in the United States every year (Scallan et al., 2011a, 2011b).

Despite the disease burden and threat, the rapid and sensitive identification of pathogens in food continues to be a challenge for those concerned with food safety. Classical microbiological methods to detect the causative agent in foodborne illnesses are laborious and often require multiple selective enrichments of the sample to achieve a presumptive identification of pathogens (Andrews et al., 2014) or to determine a reasonable assumption of the pathogen type (Naravaneni and Jamil, 2005; Velusamy et al., 2010). Many pathogens cause similar signs of disease (Scallan et al., 2011a, 2011b); therefore, there can be a delay in pathogen identification because it requires pathogen-specific screening tests, such as polymerase chain reaction (PCR). This delay can translate into an increased number of infections, which can lead to more severe and long-term impacts and a decreased ability to find the source of contamination. For example, culturing *Listeria monocytogenes* can take 3 to 7 days to yield results, and testing for *Campylobacter* spp. can take 4 to 9 days to confirm a negative result and 14 to 16 days to confirm a positive result (Velusamy et al., 2010**)**. PCR technology allows for testing of multiple pathogens at once, but it still requires some prior knowledge of the sample and an enrichment step to generate a sufficient amount of pathogen nucleic acid for PCR detection (Naravaneni and Jamil, 2005; Velusamy et al., 2010**)**. The mass spectrometry proteomics method (MSPM) for pathogen identification has the potential to significantly reduce these impacts by shortening the lag period that has been experienced with the use of conventional microbiological methods.

The MSPM was developed for the identification and classification of pathogens and does not require prior knowledge of the agent in the sample or selective enrichment steps (Jabbour et al., 2010). The output of the MSPM provides a strong and effective proteomic fingerprint method that is complementary to genomic-based techniques (i.e., microarrays and

PCR). The MSPM serves as an effective and nonrestrictive screening tool for other more targeted testing and allows for PCR analysis to confirm the identified pathogens. In addition, previous studies have shown the effectiveness of MSPM for identifying virulence factors within a pathogen and for finding biomarkers that can indicate whether or not the DNA of the pathogen was altered for increased virulence, infectivity, or pathogenicity (Jabbour et al., 2010). All of these benefits can lead to more rapid detection of a pathogen, determination of its public health threat, and indication of whether or not the pathogen was engineered for malicious intent.

The purpose of this pilot study was to determine the validity of the MSPM in ascertaining whether or not a homogenous food substance is contaminated with a common foodborne pathogen. This proof of concept study will allow for decision-makers to determine whether or not to pursue this technology as a screening or diagnostic tool for food-based laboratory testing.

2. MATERIALS AND METHODS

2.1 Preparation of Bacterial Stocks

The U.S. Army Public Health Command Region-South, DoD Food Analysis and Diagnostic Laboratory (APHC FADL; Houston, TX) prepared all of the pathogen samples. APHC FADL conducts microbiological testing according to American Association of Laboratory Accreditation (Frederick, MD). Five pathogens that were identified as common causes of foodborne illness were characterized using the MSPM and were included in a small library for MSPM analysis. The pathogens were analyzed at a concentration of approximately 10^6 colony-forming units (cfu)/mL to construct the proteomic fingerprint. The five pathogens (Table 1) that were identified as common causes of foodborne illness were *Escherichia coli* O157:H7 (U.S. Department of Agriculture [USDA] strain 43895), *Salmonella enterica* serotype Newport (USDA strain 15480), *Listeria monocytogenes* (American Type Culture Collection [ATCC] 11994), *Staphylococcus aureus* (ATCC 6538), and *Bacillus cereus* (ATCC 10876) (Center for Food Safety and Applied Nutrition, 2005; Scallan et al., 2011b). Certificates of analysis for commercially purchased bacteria stocks were obtained to ensure organism purity. A quality-control assessment of each bacterial stock was performed to include the colonial morphology and key biochemical reactions that are characteristic of each strain.

Table 1. Bacterial Organism Concentrations for Mass Spectrometry Library

Organism	Strain Number	Concentration (cfu/mL)
E. coli O157:H7	43895 (USDA strain)	0.89×10^6
S. enterica serotype Newport	15480 (USDA strain)	1.0×10^6
L. monocytogenes	11994 (ATCC strain)	2.8×10^6
S. aureus	6538 (ATCC strain)	1.6×10^6
B. cereus	10876 (ATCC strain)	0.15×10^6

The five aerobic bacterial pathogens were cultured onto trypticase soy agar (Beckton, Dickinson, and Company; Franklin Lakes, NJ) with 5% sheep blood agar (SBA) from frozen stock at 37 ± 2 °C for 18–24 h. A second culture passage to SBA for each bacterial stock was incubated overnight at 37 ± 2 °C for 18–24 h to ensure purity and typical colonial morphology. Subcultures were incubated overnight at 37 °C for 18–24 h. Viable cell density (cfu/mL) for each culture was verified using a turbidometric method with a McFarland standard inoculum (Vitek Densichek; bioMerieux, Inc.; Durham, NC) and by plating serial dilutions made in trypticase soy broth (TSB). To prepare the cultures, 100 µL of selected serial dilutions of each bacterial stock were spread-plated on SBA at 37 °C for 24 h, followed by colony count verification to determine the starting bacterial concentration for each serial dilution. Serial dilutions of bacterial stocks were frozen at –80 °C then shipped to the U.S. Army Edgewood Chemical Biological Center (ECBC; Aberdeen Proving Ground, MD) for mass spectrometry analysis and creation of the MSPM library. To determine the percent recovery of viable bacteria after freezing, frozen stock of *S. aureus* was serially diluted and plated for a colony count verification of the starting bacterial concentration. The starting bacterial concentration of the frozen *S. aureus* was almost identical to the bacterial concentration before freezing. This small, five foodborne pathogen library was used to create a reference mass spectrometry database to serve as the data source for pathogen identification.

2.2 Preparation of Mashed Potato Samples Spiked with Foodborne Pathogens

Aliquots of *S. enteric* serotype Newport were spiked into mashed potato samples. First, mashed potato samples were prepared by adding sterile water to instant mashed potatoes (Hill Country Fare brand; H-E-B; San Antonio, TX) using aseptic techniques, followed by mixing to ensure a homogenous mixture. Next, 2.3 mL of a 1×10^7 cfu/mL of *S. enteric* serotype Newport bacterial suspension, which was prepared in TSB media, was spiked into 23 mL of prepared mashed potatoes. Positive spiked samples were prepared in a biological safety cabinet and well mixed to ensure homogeneity in the sample. Negative samples consisted of 25 mL of prepared instant mashed potatoes only. Cross-contamination was eliminated by preparing negative samples in a dedicated reagent hood before spiking the positive samples.

In total, 75 pairs of spiked samples and negative controls (150 total samples) allowed for an estimation of a 95% sensitivity and specificity with 95% confidence, an allowable error of 5%, and a power of 80%. Sample pairs were marked from 1 to 75, and each member of the pair was randomly marked as A or B. The identities of spiked and unspiked samples were blinded until the completion of MSPM analysis at ECBC.

All samples were stored at –70 °C and shipped overnight on dry ice, using a certified shipper, from the APHC FADL to the ECBC. Standard guidelines for food-receiving and -handling procedures were followed.

2.3 Sample-Processing Approach

Mashed potato samples were vortexed in the sample tubes that were received (25 mL sample in a 50 mL conical bottom tube). Approximately 1 mL of the mashed potato sample was pipetted into 9 mL of phosphate-buffered saline (PBS) and vortexed to suspend any

bacterial cells in solution. The 10 mL tube was centrifuged at 400 × *g* for 20 min to pellet large pieces of mashed potatoes and leave the bacterial cells in solution. The supernatant was decanted into a new 10 mL tube and centrifuged at 6600 × *g* for 20 min to pellet the bacterial cells. The supernatant was discarded, and the pellets were washed and resuspended two times with 1 mL PBS then centrifuged at 6600 × *g* for 20 min to pellet the bacterial cells again to remove contaminants. Pellets were then resuspended with 1 mL PBS for bead-beating, which disrupted the bacterial cells. The subsequent protocol for the denaturing and trypsin digestion of the proteins extracted from the mashed potato samples was performed as previously described (Velusamy et al., 2010). The resulting tryptic peptides were analyzed using a liquid chromatography–tandem mass spectrometry (LC–MS/MS) technique.

2.4 LC–MS/MS Analysis of Tryptic Peptides

The tryptic peptides were separated using a capillary Hypersil C18 column (300 Å, 5 μm, 0.1 mm i.d. × 100 mm) with the Ultimate 3000 from Thermo Fisher Scientific (Waltham, MA). The elution was performed using a linear gradient from 98% aqueous phase (A) (0.1% formic acid [FA]) and 2% organic phase (B) (0.1% FA in acetonitrile) to 60% B over 60 min at a flow rate of 200 μL/min, which was followed by 20 min of isocratic elution. The separated peptides were electrosprayed into a linear ion trap quadrupole mass spectrometer (LTQ-XL; Thermo Fisher Scientific) at a flow rate of 0.2 μL/min. Product ion mass spectra were obtained in the data-dependent acquisition mode that consisted of a survey scan over the mass-to-charge ratio (*m/z*) range of 400–2000, followed by seven scans on the most intense precursor ions that were activated for 30 ms by an excitation energy level of 35%. A dynamic exclusion was activated for 3 min after the first mass spectrometry/mass spectrometry (MS/MS) spectrum acquisition for a given ion. Uninterpreted product ion mass spectra were searched against a microbial database with TurboSEQUEST software (Bioworks 3.1, Thermo Fisher Scientific) followed by application of an in-house proteomic algorithm for bacterial identification.

2.5 Protein Database and Database Search Engine

A protein database was constructed in a FASTA format using the annotated bacterial proteome sequences that were derived from fully sequenced chromosomes of all available *E. coli* O157:H7 (USDA strain 43895), *S. enterica* serotype Newport (USDA strain 15480), *L. monocytogenes* (ATCC 11994), *S. aureus* (ATCC 6538), and *B. cereus* (ATCC 10876) strains and more than 120 common laboratory contaminant proteins. We used the PERL program (Active State, 2011) to download these sequences automatically from the National Institutes of Health, National Center for Biotechnology website (2015). Each database entry for a given protein sequence has information about a source organism and about a genomic position of the respective open reading frame embedded into a header line. The constructed bacterial proteome database resulted from translating putative protein-coding genes and consisted of the in silico digested proteins, using trypsin and their corresponding tryptic peptides amino acids sequences. We used SEQUEST (Eng et al., 1994) to generate the in silico tryptic peptides, and two missed cleavages were allowed during this process.

The experimental MS/MS spectral database of bacterial peptides was searched using the SEQUEST (Eng et al., 1994) algorithm against the constructed proteome database of microorganisms. The SEQUEST thresholds for searching the product ion mass spectra of peptides were correlation score (Xcorr), relative correlation score (ΔCn), specificity (Sp), relative specificity (RSp), and change in the mass of the peptide (ΔMpep). The top peptide hits generated by SEQUEST were filtered with a $\Delta Cn > 0.1$, and the filtered hits were accepted as peptide identifications when their Xcorrs were higher than the thresholds that allowed the generation of a desired false discovery rate value (Peng et al., 2003).

The identification and classification of the bacterial pathogens in the analyzed samples were performed using an algorithm, developed and patented in-house, known as agents of biological origin identification (ABOid) (Deshpande et al., 2011). The ABOid algorithm process transformed the SEQUEST results, which were obtained by searching the product ion mass spectra of peptide ions against the constructed proteome database, into a taxonomically meaningful and easy to interpret output. Each selected peptide was verified for its true positive assignment using the PeptideProphet algorithm (Keller et al., 2002). The validated peptides were populated in a sequence-to-bacterium binary matrix of assignments (Deshpande et al., 2011**)**. Validated peptide sequences with a probability score of 95% and higher were retained, and each of those peptides were matched for their presence against each bacterial or laboratory contaminant in the constructed proteome database. The resulting binary bitmap was translated into a histogram output that reflected the number of matches for a given bacteria in the database. Furthermore, we used phylogenetic relationships among all strains in the constructed bacterial database as part of the decision tree process. A protein was identified as present when it was matched with at least two or more validated peptides in an analyzed sample. The ABOid algorithm inferred identification of the analyzed sample using assignments of organisms to taxonomic groups (phylogenetic classification). This assignment was based on a taxonomic hierarchy that began classification at the phylum level and followed through classes, orders, families, genus, species, and then down to the strain level.

2.6 Cluster Analysis

The output file generated by the database-searching tool COMET (Eng et al., 2013) was submitted to the ABOid algorithm, which took into consideration parameters such as sample number, spectral number, charge state of each spectra, retention time, Xcorr, RSp, SP, mass plus hydrogen (M+H), peptide, accession number, and PeptideProphet score (Scallan et al., 2011b) for the identification of the microbe in the given sample.

All samples and their corresponding identified protein accession numbers were used to generate a matrix of 144 samples (column) × 17,890 proteins (rows). For a given sample, a protein match with a bacterial protein in the database was given a score of “1” and no match was given a score of “0”. This binary matrix was then used to generate the cluster analysis using the Ward’s method for amalgamation rule and the Euclidean degree of similarity distancing rules.

2.7 MSPM Results Analysis

A third party from the Armed Forces Health Surveillance Center (Silver Spring, MD) collected all results for analysis. The diagnostic sensitivity and specificity of the MSPM for detection of *S. enterica* serotype Newport in the mashed potatoes was calculated, along with their 95% confidence intervals (CIs). Microsoft Excel software (Microsoft Corporation; Redmond, WA) was used for this statistical analysis.

3. RESULTS

The MSPM was used to determine the pathogen type by comparing the number of unique peptides identified in the sample to the theoretical peptide fingerprint in the proteomic library (Table 2). Table 3 demonstrates that for each of the paired mashed potato samples, the MSPM was used to identify the contaminated member of the pair and to detect and identify the pathogen present in all of the analyzed samples.

Table 2. Comparison of the Experimental Pathogen Samples to Their Theoretically Matched Pathogens Using ABOid

Laboratory Sample	Culture Concentration (cfu/mL)	ABOid Assigned Pathogen					Total Unique Peptide
		B. cereus	*E. coli*	*L. monocytogenes*	*S. enterica*	*S. aureus*	
B. cereus	0.2×10^6	92	2	7	2	2	105
E. coli	0.9×10^6	2	95	2	0	18	117
L. monocytogenes	2.8×10^6	3	1	44	2	1	51
S. enterica serotype Newport	1.0×10^6	3	18	3	68	0	92
S. aureus	1.6×10^6	1	0	3	1	58	63

Note: Gray shading is provided for clarity.

Three of the pathogen mashed potato sample tubes cracked in transit to ECBC, and the pairs were discarded from the statistical analysis; however, all the remaining samples were analyzed and processed using the MSPM. Therefore, 72 pairs of mashed potato samples were assessed for statistical evaluation of the MSPM performance (Table 3). Of these 72 pairs, all but one negative-control sample was categorized correctly, which resulted in a sensitivity that approached 100% and a specificity of 98.6% (95% CI: 95.5, 100). The overall test validity, using the ABOid findings, was 99.3% (95% CI: 97.9, 100).

Table 3. Validity Statistics of the MSPM in Detecting *S. enterica* Serotype Newport in Mashed Potato Samples

Result	***S. enterica* Serotype Newport**		
	MSPM +	**MSPM –**	**Total**
True positives	72	0	72
True negatives	1	71	72
Total	73	71	144
Standard Error (% CI)	–	–	95
Sensitivity (%)	100.0	–	–
Specificity (%)	98.6	1.4	(95.9, 100)
Overall validity (%)	99.3	0.7	(97.9, 100)

–, not applicable.

Figure 1 shows the cluster analysis classification of all of the unknown mashed potato samples that were analyzed using MSPM. This figure identifies two distinct clusters with no overlap, as indicated by the 100% separation value on the *x* axis. Closer analysis showed that all mashed potato samples that were positive for pathogen identification were found in Cluster 1, whereas all mashed potato samples that were negative for pathogen identification were found in Cluster 2.

An initial analysis of one pathogen sample out of the 144 blinded mashed potato pathogen samples had an inconclusive identification using the ABOid algorithm. A cluster analysis was used to compare this inconclusive sample to the two sets of conserved peptide clusters, which were clusters of positive and negative mashed potato samples (Clusters 1 and 2, respectively). Both of these clusters were determined from the blinded mashed potato paired samples using the MSPM process. Evaluation of this additional cluster analysis determined that the inconclusive sample had a higher correlation with proteins in the negative-control samples than with the positive-control samples from the mashed potato paired samples. As a result, the inconclusive sample was classified as a negative sample (Figure 1).

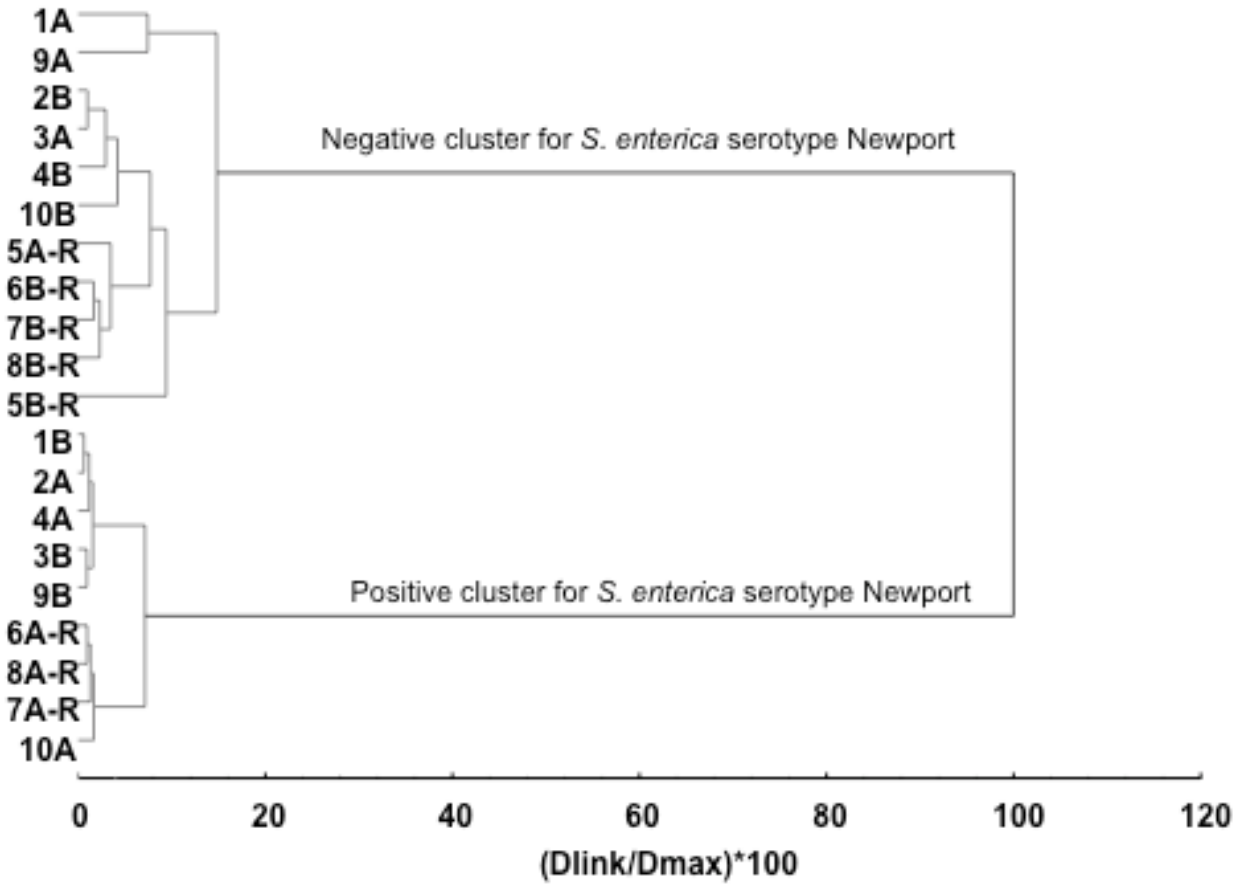

Figure 1. Classification of sample 5B based on the cluster analysis of positive and negative mashed potato samples.

Furthermore, a comparison of the protein sets from the positive mashed potato samples for *S. enterica* serotype Newport and the theoretical sets from the library sample of *S. enterica* serotype Newport in TSB was performed and is shown in Figure 2. This comparison was performed to determine the impact of sample processing on the identification process using the ABOid algorithm (Table 4). Samples of *S. enterica* serotype Newport in TSB at the same bacterial concentration as that of the contaminated mashed potato samples were analyzed. There were 724 and 655 proteins identified in the *S. enterica* serotype Newport in TSB media and the contaminated mashed potatoes, respectively, with 180 common proteins identified between the two matrices. A 9.5% decrease in the number of proteins was observed in the mashed potato samples as compared with that in the TSB media, which could be attributed to loss of bacterial proteins during sample processing of the mashed potato samples.

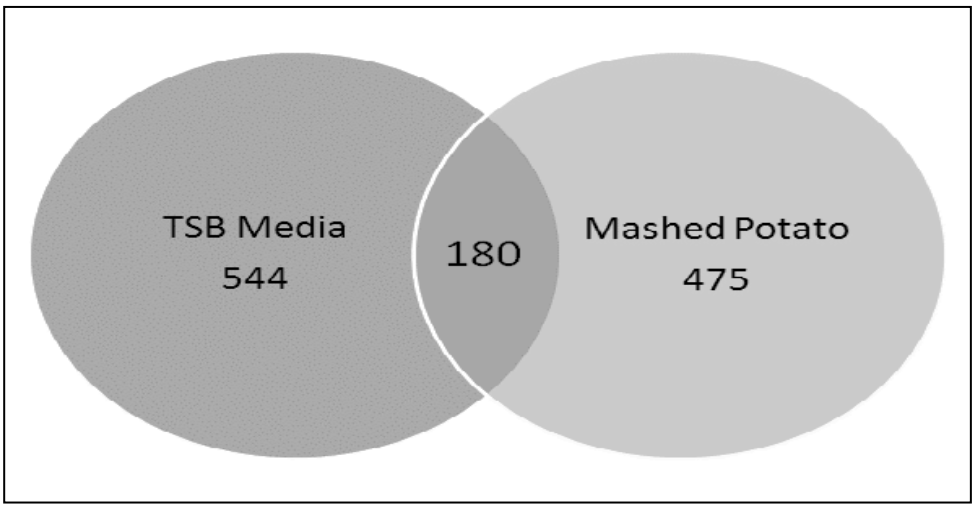

Figure 2. Comparison of the MSPM-identified proteins in mashed potato samples spiked with *S. enterica* serotype Newport versus *S. enterica* serotype Newport grown in TSB culture broth.

Table 4. A Comparison of Unique *S. enterica* Serotype Newport Peptides Identified in Spiked TSB and Mashed Potato Samples

Matrix	**TSB Media No. (%)**	**Mashed Potato No. (%)**
Total Proteins	724	655
Unique Proteins	544 (75)	475 (72.5)
Shared Proteins	180 (24.9)	180 (27.5)

In this pilot study, we also attempted to discover protein biomarkers for *S. enterica* serotype Newport in mashed potatoes that could be used for rapid screening of this organism in mashed potato and other food samples. There are 32 proteins that were identified in 38 out of 72 (53%) of mashed potato samples that were positive for *S.enterica* serotype Newport (Table 5). There were 9 proteins that were commonly identified in at least 90% or more of the 72 positive samples. Of these 9 commonly occurring proteins, osmotically inducible protein Y is a potential protein biomarker for detecting *S. enterica* serotype Newport because it is the only protein that is unique to *S. enterica* serotype Newport and was found in >95.8% of the samples analyzed. (UniProt consortium database results found at the following website: http://www.uniprot.org/uniprot/?query=salmonella+enter&sort=score [accessed 02 June 2017]; Table 4).

Table 5. Commonly Identified Proteins in *S. enterica* Serotype Newport Positive Mashed Potato Samples

Accession Number	Protein	No. of Samples Containing Protein	Occurrence (%)
YP_002043589.1	10 kDa Chaperonin (GroES protein)	69	95.83
YP_002041222.1	Flagellin	69	95.83
YP_002043404.1	50S Ribosomal protein L7\L12	69	95.83
YP_002043802.1	Osmotically inducible protein Y*	69	95.83
YP_002043590.1	60 kDa Chaperonin (GroEL protein)	68	94.44
YP_002042197.1	Phosphopyruvate hydratase	67	93.06
YP_002043697.1	Endoribonuclease L-PSP	67	93.06
YP_002040547.1	Glyceraldehyde-3-phosphate dehydrogenase	66	91.67
YP_002042751.1	Phosphoenolpyruvate carboxykinase*	65	90.28
YP_002039643.1	Peroxiredoxin-2	63	87.5
YP_002043644.1	30S Ribosomal protein S6	63	87.5
YP_002040068.1	DNA starvation\stationary phase protection protein Dps	60	83.33
YP_002042614.1	Malate dehydrogenase	58	80.56
YP_002043491.1	Stress-response protein	55	76.39
YP_002040273.1	Outer membrane protein A	54	75
YP_002040900.1	Universal stress protein F	54	75
YP_002041817.1	Serine hydroxymethyltransferase	53	73.61
YP_002043157.1	UDP-*N*-acetylglucosamine 2-epimerase	53	73.61
YP_002039690.1	Trigger factor	50	69.44
YP_002043647.1	50S Ribosomal protein L9	50	69.44
YP_002042323.1	Phosphoglycerate kinase	49	68.06
YP_002040633.1	Pyruvate kinase	48	66.67
YP_002041828.1	Phosphoribosylformylglycinamidine synthase	48	66.67
YP_002043350.1	Cell division protein ZapB	47	65.28
YP_002040008.1	Phosphoglyceromutase	45	62.5
YP_002043671.1	Inorganic pyrophosphatase	45	62.5
YP_002039240.1	Chaperone protein DnaK (HSP70) (Heat shock 70 kDa protein)*	44	61.11
YP_002039975.1	Succinyl-CoA synthetase subunit beta	43	59.72
YP_002039385.1	Dihydrolipoamide acetyltransferase	42	58.33
YP_002039854.2	Universal stress protein G	42	58.33
YP_002040983.1	YciE	40	55.56
YP_002043418.1	Transcriptional regulator HU subunit alpha	38	52.78

*Indicates potential unique protein to *S. enterica* serotype Newport.
L-PSP, liver perchloric acid-soluble protein; Dps, DNA-binding proteins from starved cells; UDP, uridine diphosphate.

4. DISCUSSION

Intentional or accidental food contamination results in a large disease burden among the U.S. population and is a threat to military readiness. There are many methods available to detect pathogens in foodstuffs, but none are rapid, unaffected by the pathogen, or free from the need for selective enrichment. The MSPM provides a new technology that can potentially be used for the detection of pathogens in food, does not require complex enrichment steps, and can return results to investigators in a short period of time.

The MSPM was accessed for its ability to be used to detect the pathogen in blinded paired samples of a homogenous food substance, with one member of the pair as the positive control and one member as the negative control. MSPM was used to correctly classify all of the positive samples and all but one negative sample. In this initial proof of concept study, a high concentration of pathogen was used in the sample to successfully show that the MSPM could detect pathogen within this high-starch food matrix. In addition, the MSPM approach provided the list of candidate proteins that can be used as biomarkers for *S. enterica* serotype Newport identification (Table 5). Although some of the most commonly occurring proteins could be found in other strains, it is noteworthy to mention that a set of peptides were strain-unique to *S. enterica* serotype Newport, and these peptides were found in at least 62 out of the 72 true positive mashed potato samples. These strain-unique peptides were associated with an osmotically inducible protein, osmY (YP_002043590.1), and peroxiredoxin-2 (YP_002039643.1). The biomarkers for these peptides can be used to develop a targeted approach to identify *S. enterica* serotype Newport, and therefore, enhance the discrimination power of MSPM to provide a rapid screening tool for *S. enterica* serotype Newport in food matrices.

In addition, MSPM was beneficial in the validation of the initial classification results. When cluster analysis of the conserved peptides shared by the pathogen was performed, the incorrectly identified positive sample was reclassified correctly as negative. This cluster analysis technique does allow for validation of the initial screening results through further statistical analysis, rather than by further laboratory analysis. This could save resources and time required to confirm the results by other conventional microbiological means, such as by culture or PCR.

5. CONCLUSION

This pilot study had a limited scope due to limited funding. Further experiments are necessary; however, the results suggest that MSPM could be a potential new technology to assist in food pathogen detection and quantification. Using MSPM allows for the identification of pathogens in mashed potato samples and for the validation of such findings through cluster analysis and taxonomic classification, without requiring multiple laboratory techniques. This technology could allow for a more rapid food pathogen detection capability, which is needed and desired by the larger public health and food safety arena. Further studies using the MSPM for identification of other microbial agents and toxins will be investigated to provide a global validation on its applicability as an emerging technology in food defense. Additional studies will

be pursued to ensure the limit of detection (LOD) statistical validity as well as detection using MSPM at concentrations near or at the LOD in mashed potato samples. Determining the LOD in this and other food matrices is critical to future research.

This study showed that the effect of the background matrix could be an issue in which a change from a relatively simple matrix (TSB media) into a mashed potato matrix resulted in a decrease of almost 10% in the number of the identified proteins (Figure 2). This factor is a challenge that will be manifested when attempting to recover pathogens from more complex food matrices. The effectiveness of the MSPM will depend on the development of effective sample preparation methods that can ensure a high recovery rate of the pathogens present within a myriad of interfering food proteins.

LITERATURE CITED

ActivePerl Software; Active State: Vancouver, BC, Canada; http://www.activestate.com/Products/ActivePerl (accessed April 2011).

Andrews, W.H.; Jacobson, A.; Hammack, T. Salmonella. In *Bacteriological Analytical Method*, 8th ed.; Revision A; U.S. Food and Drug Administration: Washington, DC, 2014.

Centers for Disease Control and Prevention. Nicotine Poisoning After Ingestion of Contaminated Ground Beef—Michigan. *Morbidity and Mortality Weekly Report* **2003**, *52* (18), 413–416.

Center for Food Safety and Applied Nutrition. *The Bad Bug Book: Handbook of Foodborne Pathogenic Microorganisms and Natural Toxins*, 2nd ed.; Food and Drug Administration: College Park, MD, 2005.

Deshpande, S.V.; Jabbour, R.E.; Snyder, A.P.; Stanford, M.; Wick, C.H.; Zulich, A.W. A Software for Automated Identification and Phyloproteomics Classification of Tandem Mass Spectrometry Data. *J. Chromatogr. Sep. Tech.* **2011**, *5* (1), 2157–7064.

Eng, J.K.; McCormack, A.L.; Yates, J.R. An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *J. Amer. Soc. Mass. Spec.* **1994,** *5* (11), 976–989.

Eng, J.K.; Jahan, T.A.; Hoopmann, M.R. Comet: An Open-Source MS/MS Sequence Database Search Tool. *Proteomics* **2013,** *13* (1), 22–24.

Food and Drug Administration. Food Protection Plan: *An Integrated Strategy for Protecting the Nation's Food Supply*; Department of Health and Human Services: Washington, DC, 2007.

Jabbour, R.E.; Deshpande, S.V.; Wade, M.M.; Stanford, M.F.; Wick, C.H.; Zulich, A.W.; Skowronski, E.W.; Snyder, A.P. Double-Blind Characterization of Non-Genome-Sequenced Bacteria by Mass Spectrometry-Based Proteomics. *Appl. Environ. Microbiol.* **2010,** *76* (11), 3637–3644.

Keller, A.; Nesvizhskii, A.I.; Kolker, E.; Aebersold, R. Empirical Statistical Model to Estimate the Accuracy of Peptide Identifications Made by MS/MS and Database Search. *Anal. Chem.* **2002,** *74* (20), 5383–5392.

Kolavic, S.A.; Kimura, A.; Simons, S.L.; Slutsker, L.; Barth, S.; Haley, C.E. An Outbreak of Shigella dysenteriae Type 2 among Laboratory Workers Due to Intentional Food Contamination. *J. Amer. Med. Assoc.* **1997,** *278* (5), 396–398.

Naravaneni, R.; Jamil, K. Rapid Detection of Food-Borne Pathogens by Using Molecular Techniques. *J. Med. Microbiol.* **2005,** *54* (1), 51–54.

National Institutes of Health, National Center for Biotechnology: Bethesda, MD; http://www.ncbi.nlm.nih.gov (accessed August 2015).

Painter, J.A.; Hoekstra, R.M.; Ayers, T.; Tauxe, R.V.; Braden, C.R.; Angulo, F.J.; Griffin, P.M. Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by Using Outbreak Data, United States, 1998–2008. *Emerg. Infect. Dis.* **2013,** *19* (3), 407–415.

Peng, J.; Elias, J.E.; Thoreen, C.C.; Licklider, L.J.; Gygi, S.P. Evaluation of Multidimensional Chromatography Coupled with Tandem Mass Spectrometry (LC/LC–MS/MS) for Large-Scale Protein Analysis: The Yeast Proteome. *J. Proteome Res.* **2003,** *2* (1), 43–50.

Scallan, E.; Griffin, P.M.; Angulo, F.J.; Tauxe, R.V.; Hoekstra, R.M. Foodborne Illness Acquired in the United States—Unspecified Agents. *Emerg. Infect. Dis.* **2011a,** *17* (1), 16–22.

Scallan, E.; Hoekstra, R.M.; Angulo, F.J.; Tauxe, R.V.; Widdowson, M.A.; Roy, S.L.; Jones, J.L.; Griffin, P.M. Foodborne Illness Acquired in the United States—Major Pathogens. *Emerg. Infect. Dis.* **2011b,** *17* (1), 7–15.

Török, T.J.; Tauxe, R.V.; Wise, R.P.; Livengood, J.R; Sokolow, R.; Mauvais, S.; Birkness, K.A.; Skeels, M.R.; Horan, J.M.; Foster, L.R. A Large Community Outbreak of Salmonellosis Caused by Intentional Contamination of Restaurant Salad Bars. *JAMA* **1997,** *278* (5), 389–395.

Velusamy, V.; Arshak, K.; Korostynska, O.; Oliwa, K.; Adley, C. An Overview of Foodborne Pathogen Detection: In the Perspective of Biosensors. *Biotechnol. Adv.* **2010,** *28* (2), 232–254.

ACRONYMS AND ABBREVIATIONS

ΔCn	relative correlation score
ΔMpep	change in mass of the peptide
ABOid	agents of biological origin identification
APHC FADL	U.S. Army Public Health Command Region-South, DoD Food Analysis and Diagnostic Laboratory
ATCC	American Type Culture Collection
cfu	colony-forming units
CI	confidence interval
ECBC	U.S. Army Edgewood Chemical Biological Center
FA	formic acid
LC–MS/MS	liquid chromatography–mass spectrometry/mass spectrometry
LOD	limit of detection
MS/MS	mass spectrometry/mass spectrometry
MSPM	mass spectrometry proteomics method
m/z	mass-to-charge ratio
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
RSp	relative specificity
SBA	sheep blood agar
Sp	specificity
TSB	trypticase soy broth
USDA	U.S. Department of Agriculture
Xcorr	correlation score

DISTRIBUTION LIST

The following individuals and organizations were provided with one Adobe portable document format (pdf) electronic version of this report:

U.S. Army Edgewood Chemical Biological Center (ECBC), Detection Spectrometry Branch
RDCB-DRI-D
ATTN: Jabbour, R.
Wade, M.

Defense Threat Reduction Agency
J9-CBS
ATTN: Graziano, A.

Department of Homeland Security
RDCB-PI-CSAC
ATTN: Negron, A.
DHS-S&T-RDP-CSAC
ATTN: Strang, P.

Defense Technical Information Center
ATTN: DTIC OA

G-3 History Office
U.S. Army RDECOM
ATTN: Smart, J.

ECBC Technical Library
RDCB-DRB-BL
ATTN: Foppiano, S.
Stein, J.

Office of the Chief Counsel
AMSRD-CC
ATTN: Upchurch, V.

ECBC Rock Island
RDCB-DES
ATTN: Lee, K.
RDCB-DEM
ATTN: Grodecki, J.


EDGEWOOD CHEMICAL BIOLOGICAL CENTER
A U.S. Army RDECOM Laboratory