We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

9

Comparative Proteomics of Tandem Mass Spectrometry Analyses for Bacterial Strains Identification and Differentiation

Rabih E. Jabbour[1], Mary M. Wade[1], Samir V. Deshpande[2],
Michael F. Stanford[1], Alan W. Zulich[1] and A. Peter Snyder[1]
[1]U.S. Army Edgewood Chemical Biological Center,
Aberdeen Proving Ground,
[2]Science and Technology Corporation,
Edgewood, MD,
USA

1. Introduction

Bacterial proteome represents the collection of functional and structural proteins that are present in the cell. The bacterial proteome consists of diverse classes of proteins with different cellular functions. Overall, the protein content of the cell represents the majority of the cell dry weight, which makes it an ideal cellular component to be utilized for bacterial characterization (Loferer-Krobacher et al., 1998). The diversity of the bacterial proteome requires the determination, identification, and characterization of its protein content in order to understand their cellular functions (Costas et al., 1990). Moreover, studying the bacterial proteome is essential to identify pathological proteins for vaccine development, diagnose and provide counter measures to infectious diseases, and to the understanding of biological systems. The availability of microbial genomic sequencing information has led to an expansive area of researching bacterial proteomics. Proteomics studies allow addressing the functional proteins produced by the changes of genetic expressions. Using comparative proteomic studies allows the examination of bacterial strain differences, both phenotypic and genetic, bacterial growth under various nutrient and environmental conditions, i.e. nutrient type, growth phase, temperature, chemical compounds, such as antibiotics. Comparative Proteomics also provides the researcher with a tool to begin characterizing the functions of the vast proportion of "hypothetical" or "unknown" proteins elucidated from genome sequencing and database comparisons.

Comparative proteomics has been widely applied to microbial identification and characterization studies through the utilization of several mass spectrometry techniques, with tandem mass spectrometry techniques proving to be effective and reliable approach [Aebersold,2003; Anhalt & Fenselau, 1975; Dworzanski, 2006; Hillkamp,2000; Jabbour, 2005, Krishnamurthy, 2000). This chapter will address the utilization of comparative proteomics and the application of tandem mass spectrometry in the identification and differentiation of bacterial strains.

2. Overview of the utilization of tandem mass spectrometry in bacterial identification and differentiation

Mass Spectrometry techniques have been extensively used for rapid identification and differentiation of microbes in general and bacteria in particular. The most predominant mass spectrometry techniques that have been utilized for bacterial identification and differentiation include electrospray ionization tandem mass spectrometry/mass spectrometry (ESI-MS/MS); matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS); surface-enhanced laser desorption/ionization (SELDI) mass spectrometry; one- or two-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (1- or 2-D SDS-PAGE); and hybrid techniques such as combination of mass spectrometry, gel electrophoresis, and bioinformatics. Those mass spectrometry methods provide either fingerprints of the bacterial proteins, i.e. MALDI-TOF-MS technique, or amino acid sequences, from tandem MS/MS analysis, of proteins from collision-induced dissociation (CID), Electron transfer dissociation, or post-source decay (PSD) of ionized tryptic peptides derived from bacterial proteins, i.e. ESI-MS/MS technique. This chapter will address the utilization of tandem mass spectrometry techniques in the differentiation of bacterial strains.

Tandem mass spectrometry techniques have witnessed significant utilization and success in the interrogation of the protein component of a biological species, virus proteins, protein toxins, and bacteria for identification and characterization purposes (Demirev & Fenselau,2008a, 2008b; Dworzanski & Snyder, 2005; Ho, 2002; Ecker, 2005; Fox, 2002, 2006; Hofstadler, 2005; Lambert,2005; Nagele, 2003; Pennigton, 1997; Sampath, 2007; Wilkins, 2006; Williams, 2002). Investigations of the protein component in biological systems constitute the realm of proteomics (Nagele, 2003; Pennigton, 1997). The LC- tandem MS technique is well-suited and equipped to handle the complex and very comprehensive suites of proteins, in a reproducible fashion (William 2002), present in biological threat microorganisms. The vast amount of protein and peptide data generated from a typical LC-tandem MS analysis needs to be addressed in an efficient and timely manner. Data reduction techniques have spawned a number of successful bioinformatics software analysis tools to efficiently address this task (Fox, 2002, 2006; Yates, 1998; Kuwana, 2002). Furthermore, new genomes are constantly being realized and resolved so as to increase the database of bacterial genomes to interrogate a biological sample (Dworzanski & Snyder, 2005). A major portion of the Centers for Disease Control (CDC) Category A, B, and C biological threats have their genomes fully sequenced and available for bioinformatics coupled to MS-based proteomics (NCBI website, 2010; Integrated genomic, 2010; Rotz, 2002).

The US Government has initiated extensive efforts in the detection and identification of biological threat species in their Defense Advanced Research Projects Agency (DARPA) programs that explore the "detect to protect" and "detect to treat" paradigms (National research Council [NRC], 2005; Demirev, 2005). Those initiatives cover areas of general health risk, bio-terrorism utility, Homeland Security, agricultural monitoring, quality of foodstuffs, environmental monitoring, and biological warfare agents in battlefield situations (Demirev & Fenselau,2008a). Some of the concerns include incidents such as a ricin attack (Bevilacqua, 2010) and the *Bacillus anthracis* spore attack on the US postal system in the fall of 2001 (Demirev & Fenselau,2008b; Dworzanski & Snyder, 2005; Friess, 2010; Ho, 2002; Wilkins, 2006).

Proteomic analyses by LC-MS have been used in the characterization of bacteria (Castanha,2006; Dworzanski, 2004, 2006;Lambert, 2005;). Given the degree of success for tandem MS-based proteomics in bacterial characterization, a comparative proteomic study was reported about the potential of the outer membrane protein (OMP) and whole cell protein extracts, independently, can distinguish between strains of the same species (Jabbour et al., 2010). Typically, whole cell protein extracts are usually investigated or select portions of the bacterium, such as the outer membrane, are isolated and the proteins extracted there from. In the membrane, the OMPs act as active mediators between the cell and its environment and are often associated with virulence in Gram-negative pathogens. In pathogenic *Escherichia coli*, there are multiple OMPs present which are required for intestinal colonization as well as those that play a role in the type III secretion system responsible for delivering effector proteins to host cells (Garmendia, 2005; Ide,2001; McDaniel, 1997; Wachter,1999).

3. Outer membrane proteins for bacterial strains differentiation

Outer membrane proteins (OMPs) of gram-negative bacteria act as active mediators between the cell and its environment and are often associated with virulence in gram-negative pathogens (Jerse et al., 1990; Kaper et al., 2004; Koebnik et al., 2000;). Avriulent strains often lack one or more of the plasmids or genes encoding proteins needed for virulence. These differences in OMP expression between virulent and avirulent strains of gram negative bacteria could potentially be exploited to distinguish among strains. Therefore, OMPs could prove to be potential biomarkers for Bacterial strain differentiation.

The off-line 2-D chromatofocussing and reverse phase LC with electrospray-time of flight (ESI-TOF)-MS and matrix-assisted laser desorption ionization (MALDI) TOF-MS detection instrumentation have been used to analyze whole cell protein extracts of non-pathogenic and pathogenic (O157:H7) *E. coli* strains (Zheng, 2005). Those analyses provided various proteins where, in addition to commonly shared proteins, seven unique proteins were found in a non-pathogenic *E. coli* strain, and five unique proteins were found to be expressed in the pathogenic O157:H7 strain. These intracellular, non-OMP proteins were the basis for distinguishing the *E. coli* strains; however, this information was not applied to bioinformatics cross-referencing with a proteome database.

A series of Enterobacteria were investigated and cross-referenced with on-line protein databases (Pribil, 2005). OMPs were investigated by MALDI-TOF- tandem MS where microgram amounts of cells were briefly subjected to trypsin digestion on a stainless steel target plate. Four Enterobacteria were investigated and protein mass spectra were analyzed. Peptide analyses provided protein identification, and multiple assignments allowed database searches for matching to the Enterobacteria species: *E. coli*, *E. herbicola*, *E. cloacae*, and *Salmonella typhimurium*. Some of the distinguishing proteins originated in the cellular milieu and unique OMPs were identified in all four species.

Top-down proteomics and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF/TOF) tandem mass spectrometries were used to differentiate protein extracts of *E. coli* strains. Six ions found in a collection of mass spectra originated from proteins that could distinguish between pathogenic and non-pathogenic *E. coli* strains by tandem TOF mass spectrometry. A unique protein biomarker ion at m/z 7705.6 was found (putative uncharacterized YahO) in pathogenic O157:H7 and pathogenic nearest neighbor O55:H7 (infantile diarrhea) strains. Another ion at m/z 9737.5 indicative of the acid stress

chaperone-like protein: HdeA was found in the O157:H7 strain. An ion (m/z 9063.4) in the mass spectrum of non-pathogenic *E. coli* RM3061 was absent in the O157:H7 mass spectrum. Tandem TOF mass spectrometry analysis identified the peak as the HdeB acid stress chaperone-like protein which was useful in discrimination for this non-pathogenic *E. coli* strain.

In another study, the membranes of the *S. typhimurium* and *Klebsiella pneumoniae* Enterobacteria were isolated, and the proteins were extracted with subsequent 2-D electrophoresis (Fagerquist, 2010). The excised protein spots were digested with trypsin and analyzed by MALDI-TOF-MS and peptide mass fingerprinting. The masses predominately originated from OMP peptides and were searched against microorganism databases for identification purposes. Twenty-five and fourteen unique proteins were found in *S. typhimurium* and *K. pneumoniae*, respectively, in a reproducible fashion (Lamontagne, 2007). Pathogenic *E. coli*, such as the O157:H7 strain is a public health pathogen responsible for most common food borne and waterborne illnesses. This bacterium contains a full complement of OMP proteins.

Yersinia pestis is classified as a Category A pathogen and is an important potential biowarfare agent. Virulent *Y. pestis* contains three plasmids encoding multiple OMPs that are required for virulence (Ben-Gurion & Shafferman, 1981; Ferber, 1981; Filippov, 1990). For example, the pCD1 plasmid encodes several *Yersinia* OMPs and a type III secretion system, which are needed for survival and entry into host eukaryotic cells (Cornelis, 2002; Ramamurthi, 2002). Additionally, the pPCP1 plasmid encodes an OMP plasminogen activator that interferes with clotting and complements (Titball, 2003). Avirulent strains often lack one or more of the plasmids or genes encoding proteins needed for virulence, and it is these differences in OMP expression between virulent and avirulent strains of Gram-negative Enterobacteria that could potentially be exploited in order to distinguish among strains.

Alternatively, high-throughput tandem mass spectrometry-based proteomics was applied as a means for characterizing cellular proteins and producing amino acid sequence information for peptides derived from these proteins for *E. coli* and *Y. pestis*. Whole cell protein and cell membrane OMP extracts were compared and contrasted with the in-house BACid bioinformatics modeling tools for species and strain level discrimination (Jabbour, 2010).

4. Bioinformatics tools for bacterial strains differentiation using tandem mass spectrometry

Utilization of MS techniques for bacterial differentiation relies on the comparison of the proteomic information generated from either intact protein profiles (top-down) or the product ion mass spectra of digested peptide sequences (bottom-up) analyses (Warscheid, 2003; Washburn, 2001). For top-down analysis, bacterial differentiation is accomplished through the comparison of the MS data of intact proteins with an experimental mass spectral database containing the mass spectral fingerprints of the studied microorganisms (Craig, 2004). Conversely, bacterial differentiation using the product ion mass spectral data of digested peptide sequences is accomplished through the utilization of search engines against publically available sequence databases to infer identification (Eng, 1994; Warscheid 2004). Several peptide searching algorithms (i.e. SEQUEST and MASCOT) have been developed to address peptide identification using proteomics databases that were generated from either fully or partially genome sequenced organisms (Craig, 2004; Xiang, 2000).

Recent developments in the microbial differentiation field have focused on improving the selectivity of the MS data processing. The product ion mass spectrum-SEQUEST approach was reported for the identification of specific bacteria using a custom-made, limited database of sequences (Keller, 2002; VerBerkmoes, 2005). Another approach used open reading frame (ORF) translator programs to predict possible protein sequences from all probable ORFs and correlate them with the genomic sequences to establish an identification of microorganisms (Chen, 2001). This approach did not show advantages over the product ion mass spectrum method with regard to strain level discrimination (Wolters et al., 2001). However, a recent advancement in proteomics approaches to bacterial differentiation reported a hybrid approach combining protein profiling and sequence database searching using accurate mass tag (Lipton et al., 2002; Norbeck et al., 2006). This approach was used to probe defined mixtures of bacteria to evaluate its capabilities.

Alternatively, an emerging bioinformatics approach that is based on a cross correlation between the product ion spectra of the tryptic peptides and their corresponding bacterial proteins derived from an in-house comprehensive proteome database from genome sequenced microorganisms has been validated (Jabbour, 2010). The exploitation of this proteome database approach allowed for a faster search of the product ion spectra than that using genomic database searching. Also, it eliminates inconsistencies observed in publicly available protein databases due to the utilization of non-standardized gene finding programs during the process of constructing the proteome database. The proposed approach uses an ensemble of bioinformatics tools for the classification and potential identification of bacteria based on the peptide sequence information. This information is generated from the liquid chromatography tandem mass spectrometry (LC-MS-MS) analysis of tryptic digests of bacterial protein extracts and subsequent profiling of the sequenced peptides to create a matrix of sequence-to-microbe (STM) assignments. This proteomics approach is an unsupervised approach to reveal the relatedness between the analyzed samples and the database of microorganisms using a binary matrix approach. The binary matrix is analyzed using diverse visualization and multivariate statistical techniques for bacterial classification and identification.

5. Experimental methods

5.1 Bacterial strains growth and culture conditions

Pathogenic strains employed in the present study were *E. coli* O157:H7 and *Y. pestis* Colorado 92 (CO92). Non-pathogenic strains employed were *E. coli* K-12 and *Y. pestis* A1122. Working cultures were prepared by streaking cells from cryopreserved stocks onto tryptic soy agar (TSA) followed by incubation for approximately 18 hours at 37º C for *E. coli* and 30º C for *Y. pestis* strains. After incubation, all working culture plates were stored at 4º C. Cells from working cultures were used to inoculate broth cultures for each strain, which consisted of 100 mL of trypticase soy broth (TSB) for *E. coli* strains and 100 mL of brain heart infusion (BHI) for *Y. pestis* strains. Cultures were incubated for approximately 18 hours at 37º C for *E. coli* strains and 30º C for *Y. pestis* strains with rotary aeration at 180 rpm. After incubation, broth cultures were pelleted by centrifugation (2,300 RCF at 4º C for 10 min), washed, and resuspended in 10 mL HEPES buffer followed by heating at 95 ºC for 1 hour to lyse the cells. After heating, a portion of each sample was plated onto TSA and incubated for five days at the appropriate temperature to ensure no growth prior to removing samples from the BSL-2 or BSL-3 laboratory. Total cellular protein samples (whole cell protein extracts) were heated

for one hour to ensure that a no growth situation was confirmed on agar plates for safety concerns.

5.2 Isolation of the Outer Membrane Proteins (OMPs)

After lysis of the whole cells by heating at 95° C for one hour, the cell debris was pelleted by centrifugation at 2,300 RCF at   4° C for 10 min. The supernatant was then centrifuged at 100,000 x g for one hour to pellet the proteins. The pellet was resuspended in 1 mL of HEPES buffer, 1 mL of a 2% Sarkosyl solution (N-lauroylsarcosine sodium salt solution) was added, and the sample was incubated at room temperature for 30 min. Samples were centrifuged at 100,000 x g for one hour, and the pellet containing OMPs was resuspended in 1 mL of HEPES buffer.

5.3 Processing of whole cell lysates and OMPs samples

All protein samples were ultrasonicated (20 seconds pulse on, 5 seconds pulse off, and 25% amplitude for 5 min duration) and a small portion of the lysates was reserved for 1-D gel analysis. The lysates were centrifuged at 14,100 x g for 30 min to remove any debris. The supernatant was then added to a Microcon YM-3 filter unit (Millipore, Catalogue # 42404) and centrifuged at 14,100 x g for 30 min. The effluent was discarded. The filter membrane was washed with 100 mM ABC and centrifuged for 20 min at 14,100 x g. Proteins were denatured by adding 8 M urea and 3 µg/µL DTT to the filter and incubating overnight at 37° C on an orbital shaker at 60 rpm. Twenty microliters of 100% ACN was added to the tubes and allowed to incubate at room temperature for 5 min. The tubes were then centrifuged at 14,100 x g for 40 min and washed three times using 150 µL of 100 mM ABC solution. On the last wash, ABC was allowed to sit on the membrane for 20 min while shaking, followed by centrifugation at 14,100 x g for 40 min. The micron filter unit was then transferred to a new receptor tube and the proteins were digested with 5 µL trypsin in 240 µL of ABC solution + 5 µL ACN. Proteins were digested overnight at 37° C on an orbital shaker set to 55 rpm. Sixty microliters of 5% ACN/0.5% formic acid (FA) was added to each filter to quench the trypsin digestion followed by two minutes of vortexing for sample mixing. The tubes were centrifuged for 30 min at 14,100 x g. An additional 60 µl 5% ACN/0.5% FA mixture was added to the filter and centrifuged. The effluent was then analyzed using LC-ESI- tandem MS.

5.4 LC-tandem MS analysis of peptides

 The tryptic peptides were separated using a capillary Hypersil C18 column (300 Å, 5 µm, 0.1 mm i.d. × 100 mm) by using the Surveyor LC from ThermoFisher (San Jose, CA 95101). The elution was performed using a linear gradient from 98% A (0.1% FA in water) and 2% B (0.1% FA in ACN) to 60% B over 60 min at a flow rate of 200 µL/min, followed by 20 minutes of isocratic elution. The resolved peptides were electrosprayed into a linear ion trap mass spectrometer (LTQ, Thermo Scientific, San Jose, CA 95101) at a flow rate of 0.8 µL/min. Product ion mass spectra were obtained in the data dependent acquisition mode that consisted of a survey scan over the m/z range of 400-2000 followed by seven scans on the most intense precursor ions activated for 30 ms by an excitation energy level of 35%. A dynamic exclusion was activated for 3 min after the first MS-MS spectrum acquisition for a given ion. Uninterpreted product ion mass spectra were searched against a microbial database with TurboSEQUEST (Bioworks 3.1, Thermo Scientific, San Jose, CA 95101)

followed by application of an in-house proteomic algorithm for bacterial identification of the replicate analyses.

5.5 Protein database and database search engine

A protein database was constructed in a FASTA format using the annotated bacterial proteome sequences derived from fully sequenced chromosomes of 1433 bacteria, including their sequenced plasmids (as of May 2011). A PERL program (http://www.activestate.com/Products/ActivePerl) was written to automatically download these sequences from the National Institutes of Health National Center for Biotechnology (NCBI) site (http://www.ncbi.nlm.nih.gov). Each database protein sequence was supplemented with information about the source organism and genomic position of the respective open reading frame (ORF) embedded into a header line. The database of bacterial proteomes was constructed by translating putative protein-coding genes and consists of tens of millions of amino acid sequences of potential tryptic peptides obtained by the *in silico* digestion of all proteins (assuming up to two missed cleavages).

The experimental product ion mass spectra of bacterial peptides were searched using the SEQUEST (Warscheid, 2003) algorithm against a constructed proteome database of microorganisms. The SEQUEST thresholds for searching the product ion mass spectra of peptides were Xcorr, deltaCn, Sp, RSp, and deltaMpep. The search results were filtered by using Xcorr = 1.90, 2.20, and 3.75 thresholds for peptide ions of +1, +2, and +3 charges, respectively (Ma, 2009; Wu, 2003). These parameters provided a uniform matching score for all candidate peptides. The generated outfiles of these candidate peptides were then validated using the Peptide Prophet algorithm (Keller et al., 2002). Peptide sequences with a probability score of 95% and higher were retained in the dataset and used to generate a binary matrix of sequence-to-bacterium (STB) assignments. The binary matrix assignment was populated by matching the peptides with corresponding proteins in the database and assigning a score of one. A score of zero was assigned for a non-match. The column in the binary matrix represents the proteome of a given bacterium, and each row represents a tryptic peptide sequence from the LC product ion mass spectral analyses. A sample microorganism was matched with a database bacterium by the number of unique peptides that remained after filtering of degenerate peptides from the binary matrix. Verification of the classification and identification of candidate microorganisms was performed through hierarchical clustering analysis and taxonomic classification (Jabbour et al., 2010).

The SEQUEST-processed product ion mass spectra of the peptide ions were compared to an NCBI protein database with the in-house BACid developed software (Dworzanski et al., 2006). BACid provided a taxonomically meaningful and easy to interpret output. It calculated the probabilities that a peptide sequence assignment to a product ion mass spectrum was correct and used accepted spectrum–to-sequence matches to generate an STB binary matrix of assignments. Validated peptide sequences, either present or absent in various strains (STB matrices), were visualized as assignment bitmaps and analyzed by the BACid module that used phylogenetic relationships among bacterial species as part of a decision tree process. The bacterial classification and identification algorithm used assignments of organisms to taxonomic groups (phylogenetic classification) based on an organized scheme that begins at the phylum level and follows through the class, order, family, genus, and species to the strain level. BACid was developed in-house using PERL, MATLAB and Microsoft Visual Basic.

6. Results and discussion

6.1 Comparative proteomic differentiation between the whole cell and the OMP extracts for the *E. coli* O157:H7 strain

The whole cell protein extracts of *E. coli* strain O157:H7 were prepared and analyzed by LC-Tandem ESI-MS/MS. The bioinformatics analyses involved the nearest-neighbor analysis, using the Euclidean single linkage approach to arrive at a set of proteins for species and strain matching to the database.

Figure 1 shows the identification and classification of the experimental sample, whole cell extract, as *E. coli* O157:H7 strain. However, this identification is equally shared with *E. coli* UTI89, which is the causative agent of human urinary tract infections. Although *E. coli* UTI89 is related to *E. coli* O157:H7, it is missing certain proteins such as the OMP HU2 outer membrane and flagella related proteins that are distinctly expressed in *E. coli* O157:H7 (*vide infra*). A comparative proteome list of the strain-unique proteins and the total number of identified proteins for the mentioned *E. coli* O157:H7 extracts is shown in table 1. There are five and eight unique proteins resulted from the bioinformatics analysis of the peptide product ion mass spectra from the *E. coli* O157:H7 whole cell and OMPs extracts, respectively. Figure 2 shows the nearest neighbor similarity linkage results for the OMP extract of *E. coli* O157:H7. This dendogram shows an unambiguous strain level differentiation for the *E.coli* O157:H7 as compared together *E. coli* strains. It is worth mentioned that the next nearest neighbor, which is *E. coli* UT189, is relatively distant at approximately 2.2 linkage units unlike that from the whole cell protein extract (Figure 1). This result indicates that OMPs extract can potentially serves as strain-unique biomarkers for bacterial strain differentiation.

Whole Cell Extract		OMP Extract	
Accession Number	Unique Protein Name	Accession Number	Unique Protein Name
BAA35715	OMP HU2 protein	NP_310124	Acid sensitivity protein
NP_290616	50S ribosomal protein L10	NP_310689	Flagellin
NP_290256	Secreted protein EspA	NP_311482	Heat shock protein
NP_310689	Structural flagella protein	NP_308975	Hypothetical protein ECs0948
NP_312864	Two-component sensor protein	NP_309690	Outer membrane protease precursor
		NP_309226	Putative antirepressor protein
		NP_309783	Putative OMP
		NP_312404	Slp
Total Proteins	162		89

Table 1. Identified unique Proteins lists detected in the Whole Cell Protein and OMP Extracts of E. coli O157:H7.

Moreover, a closer look at the resulted bioinformatics data showed the total number of proteins identified between the two extraction techniques was such that the whole cell preparation had a significantly higher number of proteins of 162 as compared to the that of the number of OMP extract proteins of 89. However, the number of unique proteins that were identified from the OMP extract (eight proteins) was greater than that in the whole cell protein extract (five proteins) (Table 1). These numbers of unique proteins are very similar to that of the whole cell protein extracts for *E. coli* strains investigated (Zheng et al. 2005). That work found five unique proteins from the E. coli O157:H7 strain. However, this does
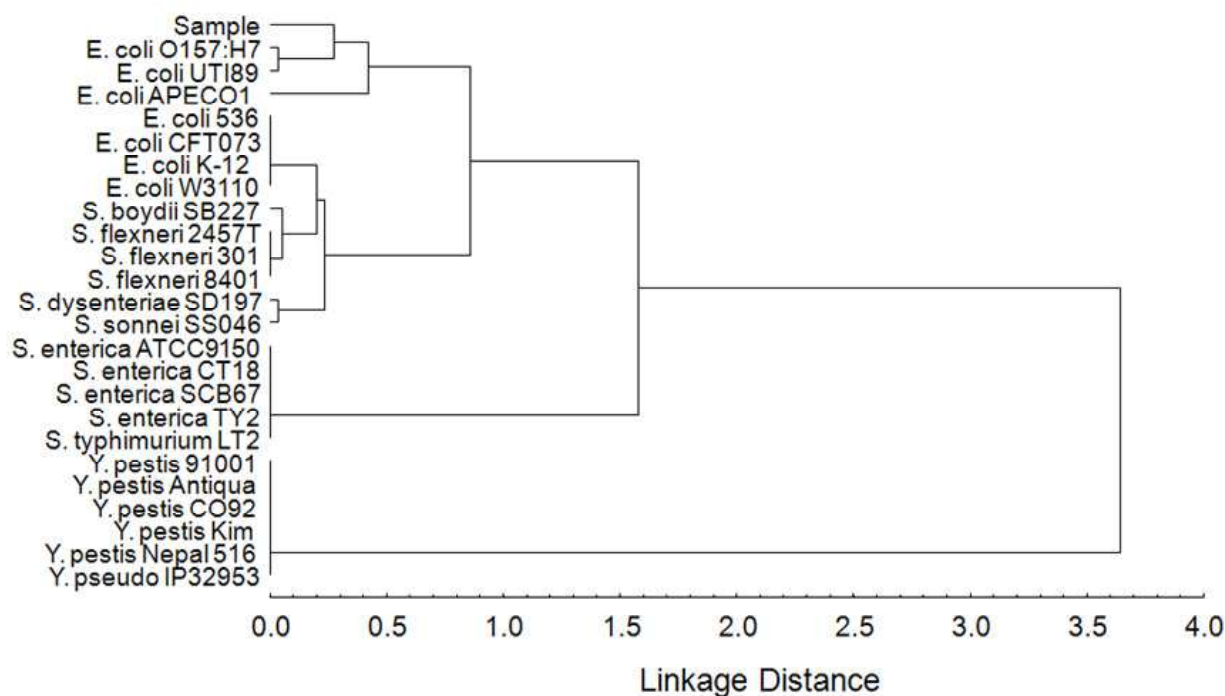
Fig. 1. Euclidean linkage similarity dendogram of the Nearest-neighbor classification of whole cell extract of E. coli O157 H:7.
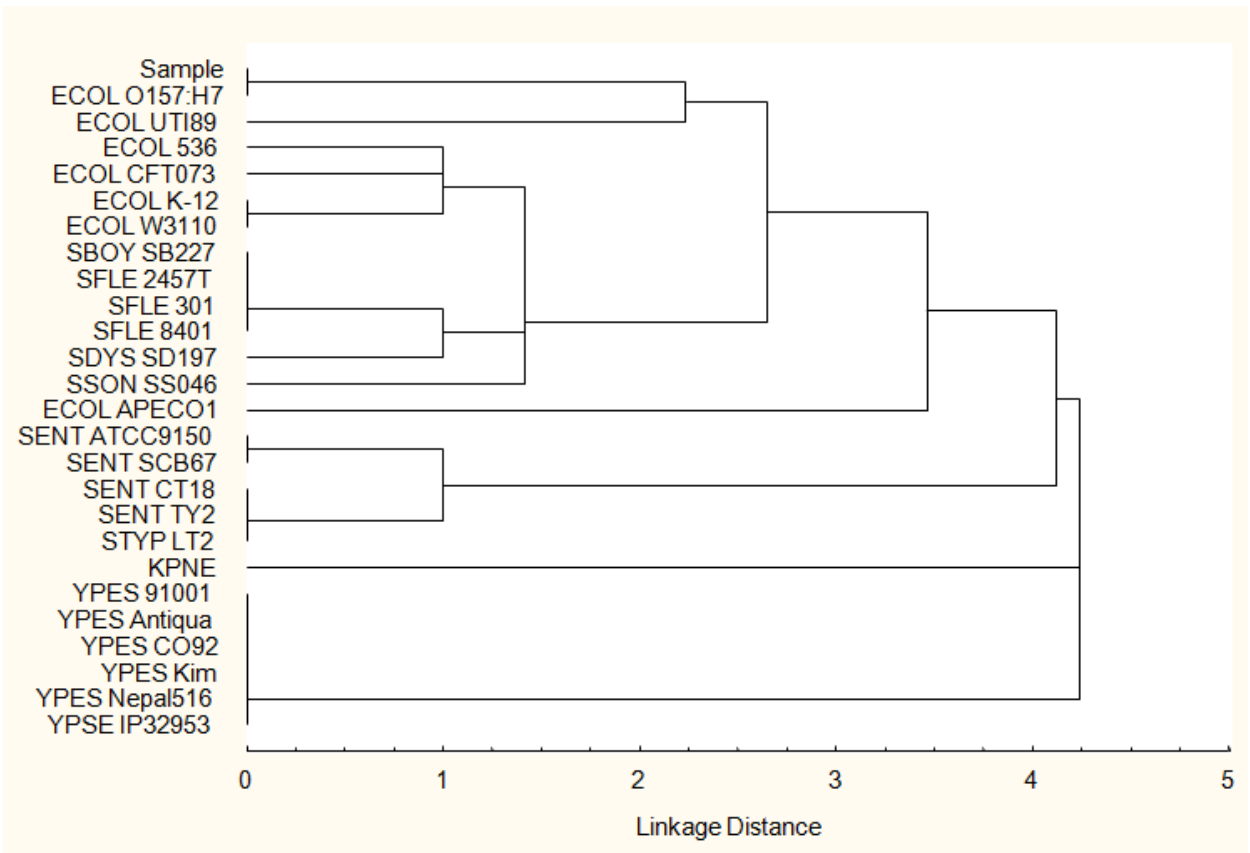

Fig. 2. Euclidean linkage similarity dendogram of the Nearest-neighbor classification of OMPs extract of *E. coli* O157 H:7.

not imply an absence of the additional OMPs in the whole cell extract. Rather it may be that a higher abundance of non-OMPs, or remaining protein in the cell, potentially suppressed the detection of the OMPs in the whole cell protein extracts by tandem MS. Mass spectral analysis can suffer from ionization suppression due to the presence of large numbers of ionizable species. Generally, a whole cell extract has a significantly larger number of ionizable peptides with a greater abundance of non-outer membrane tryptic peptides compared to that of an OMP extract. Therefore, whole cell protein extract analysis likely experiences a degree of ionization suppression during mass spectral analysis.

6.2 Comparative proteomic differentiation between the whole cell and the OMP extracts for the *E. coli* K-12 strain

The results of the bacterial strain level differentiation of the whole cell and OMPs extracts for *E. coli* K-12 are shown in Figures 3 and 4, respectively. The results indicate that those extracts provided sufficient number of identified proteins to correctly identify the *E. coli* K-12 strain. Figure 3 shows that the whole cell protein extract produced an equal similarity with the sample and the *E. coli* K-12 and W3110 strains. This is in agreement with the literature, which reported that *E. coli* W3110 is actually a substrain of K-12 (Baglioni et al., 2003; Yamada et al., 1993). It worth mentioning that the whole cell extract (Figure 3) is approximately 0.03 linkage units distant between the sample/K-12/W3110 *E. coli* group of strains and the next nearest-neighbor group that includes the *E. coli* 536/UT189/CFT73/O157:H7 strains. Hence, the whole cell protein extract was able to delineate the sample containing *E. coli* K-12 from that of the of the *E. coli* O157:H7 strain.
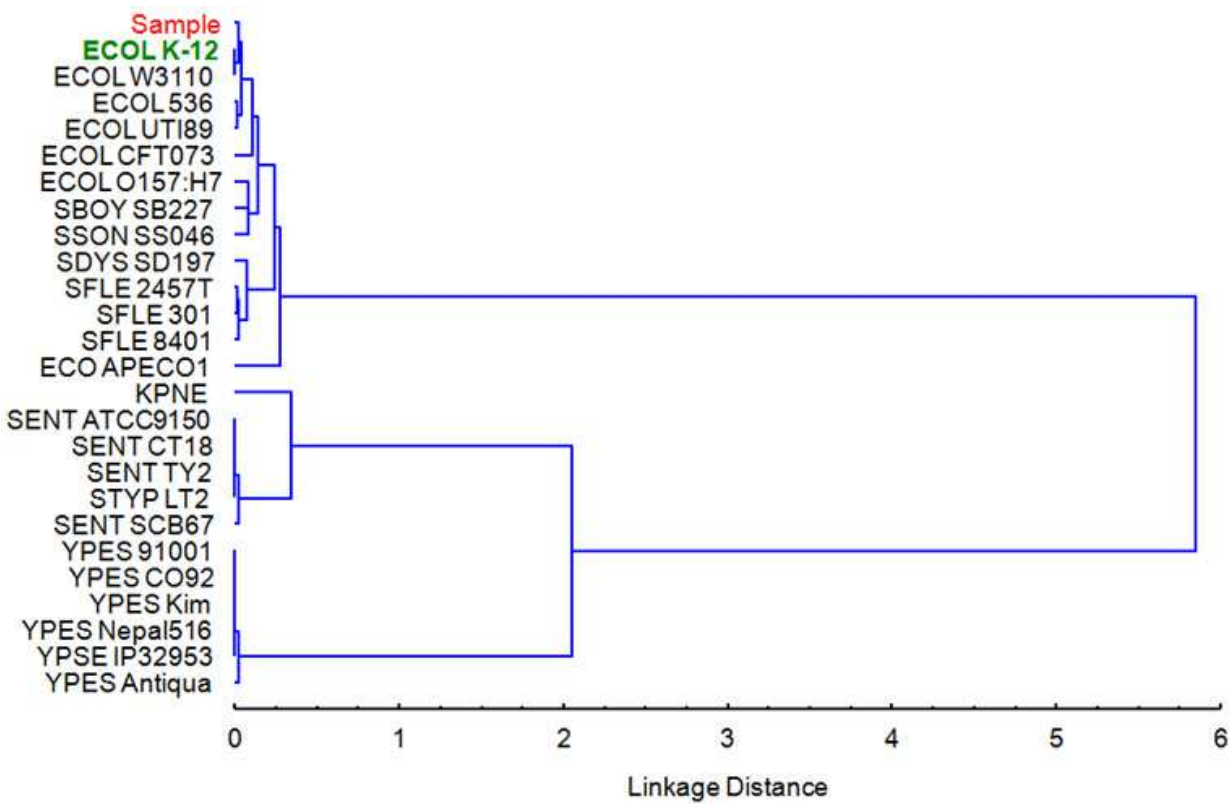

Fig. 3. Euclidean linkage similarity dendogram of the Nearest-neighbor classification of whole cell extract of E. coli K-12 strain.

Figure 4 shows the nearest neighbor Euclidean similarity linkage analysis for the OMP extracts of the *E. coli* K-12 sample. This dendogram shows that the OMP extract provided an enhancement of the strain differentiation as compared with that of whole cell extract. Although, a sample was matching with the non-pathogenic W3110 strain, however, the labels signify the same organism (*vide supra*). No ambiguity was observed in the strain differentiation. Moreover, there is a relatively larger linkage distance (0.10) between the sample/K-12/W3110 and the 536/UT189/CFT073/O157:H7 groups of *E. coli* strains from the OMP as compared to that from the whole cell extract, figure 3.
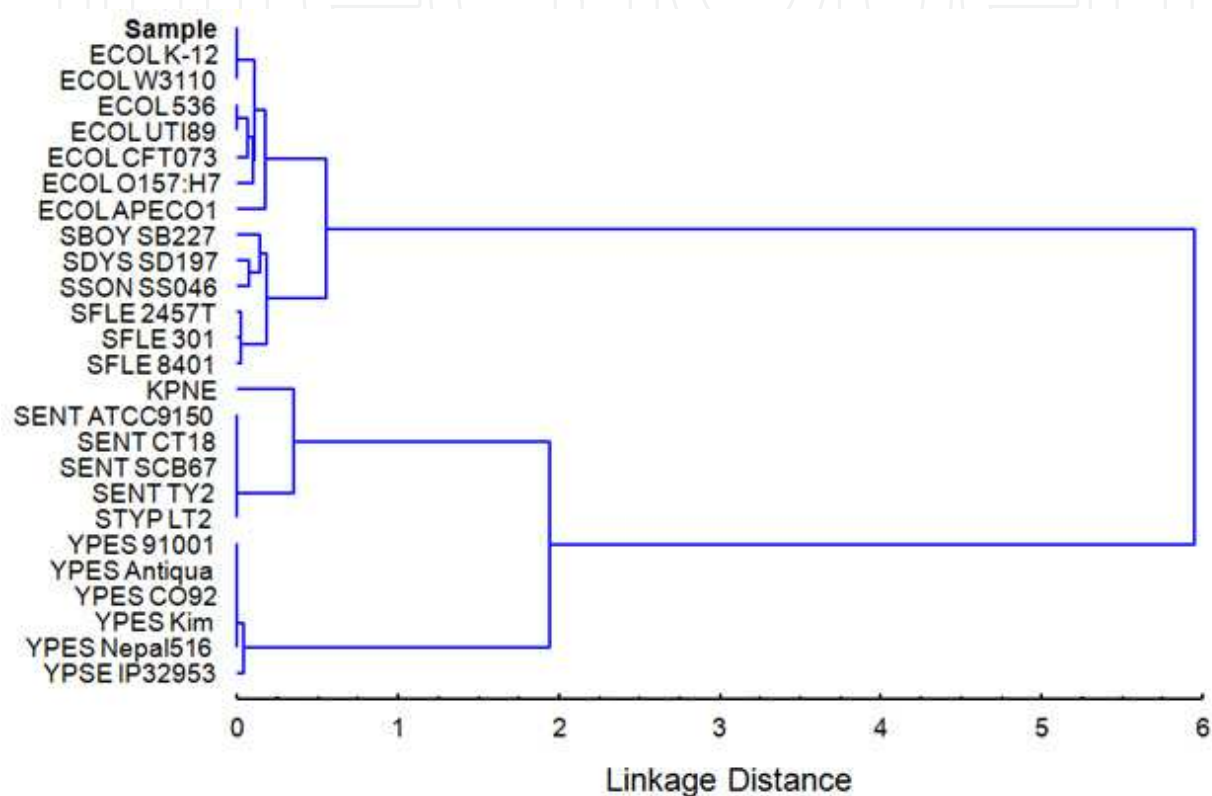

Fig. 4. Euclidean linkage similarity dendogram of the Nearest-neighbor classification of OMPs extract of E. coli K-12 strain.

Table 2 presents a list of the unique proteins. The total number of identified proteins found in the proteomics analysis for the K-12 strain was 194 and 112 for the whole cell protein and OMP extracts, respectively. The number of strain-unique proteins that were identified by the bioinformatics algorithm was greater in the OMP extracts (ten proteins) compared to that in the whole cell extracts (eight proteins). These numbers of unique proteins from the K-12 extracts are very similar to that of the whole cell protein extracts for *E. coli* strains investigated by Zheng et al. (Zheng et al., 2005). That work found seven unique proteins from the non-pathogenic *E. coli* 88-0447 (O136STa).

Overall, the comparative proteomic analyses of the E.coli whole cell extracts showed that there 162 proteins produced for *E. coli* O157:H7 strain vs. 194 for that of *E. coli* K-12 one, see tables 1-2. Upon removing the highly conserved, house-keeping, denigrate and energy transfer proteins from both strains, the number of strain-unique proteins was eight for *E. coli* K-12 and five for *E. coli* O157:H7. From analyses of the OMP protein extracts, a comparison of the total experimentally-determined number of proteins showed a difference

Whole Cell Extract		OMPs Extract	
Accession Number	**Protein Name**	**Accession Number**	**Protein Name**
YP_669714	Aspartyl-tRNA Synthetase	NP_415097	DLP12 prophage; outer membrane protease VII
NP_417795	Bacterioferrin	NP_415269	Peptidoglycan-associated outer membrane protein
NP_668903	Chorismate synthase	NP_417083	Protein disaggregation chaperone
NP_755058	GnsA/GnsB family	NP_415423	Pyruvate formate lyase I
NP_671573	Putative cytoplasmic protein	NP_415759	Oligopeptide transporter subunit
NP-415386	Lipoprotein	NP_416009	Predicted glutamate: gamma-aminobutyric acid antiporter
YP_670276	Hypothetical protein	NP_414968	Predicted lipoprotein
NP_415386	Predicted lipoprotein	NP_417320	5-keto-4-deoxyuronate isomerase
		NP_415772	OMP W
		NP_417963	Outer membrane lipoprotein
Total identified Proteins	**194**		**112**

Table 2. Identified unique proteins lists detected in the whole cell and OMP extracts of E. coli K-12 strain.

between the two *E. coli* strains. The O157:H7 strain had 89 total identified proteins compared to 112 for the K-12 strain. Upon removing the highly conserved, house-keeping, and energy transfer proteins from both strains, the number of strain-unique proteins for *E. coli* O157:H7 is eight and that for *E. coli* K-12 is ten in the OMPs extract of the studied *E. coli* strains as shown in table 2.

6.3 Comparative proteomic differentiation between the whole cell and the OMP extracts for the *Yersinia pestis* CO92 strain

A comparison of the LC-Tandem MS and bioinformatics results of the proteins present in the whole cell and OMP extracts of *Y. pestis* CO92 was performed. Figure 5 shows the identification results of the whole cell protein extract for *Y. pestis* CO92. The dendrogram indicates an ambiguous strain level differentiation between the experimental sample and the database *Y. pestis* CO92 entry.The bioinformatics analysis of the whole cell extracts of *Y. pestis* CO92 matched with five strains entries of *Yersinia* strains in the database. The CO92 experimental strain was matched to the only avirulent *Y. pestis* strain (91001) in the database as well as to the virulent Antiqua, CO92, Nepal 516, and IP32953 *Y. pestis* strains. However the Y. pestis KIM strain resided two linkage units distant from the sample and remaining five *Y. pestis* strains in the nearest neighbor similarity linkage analysis. The set of unique proteins for whole cell protein extracts of *Y. pestis* CO92 shows only four biomarkers associated with its reported virulence factors (Table 3).
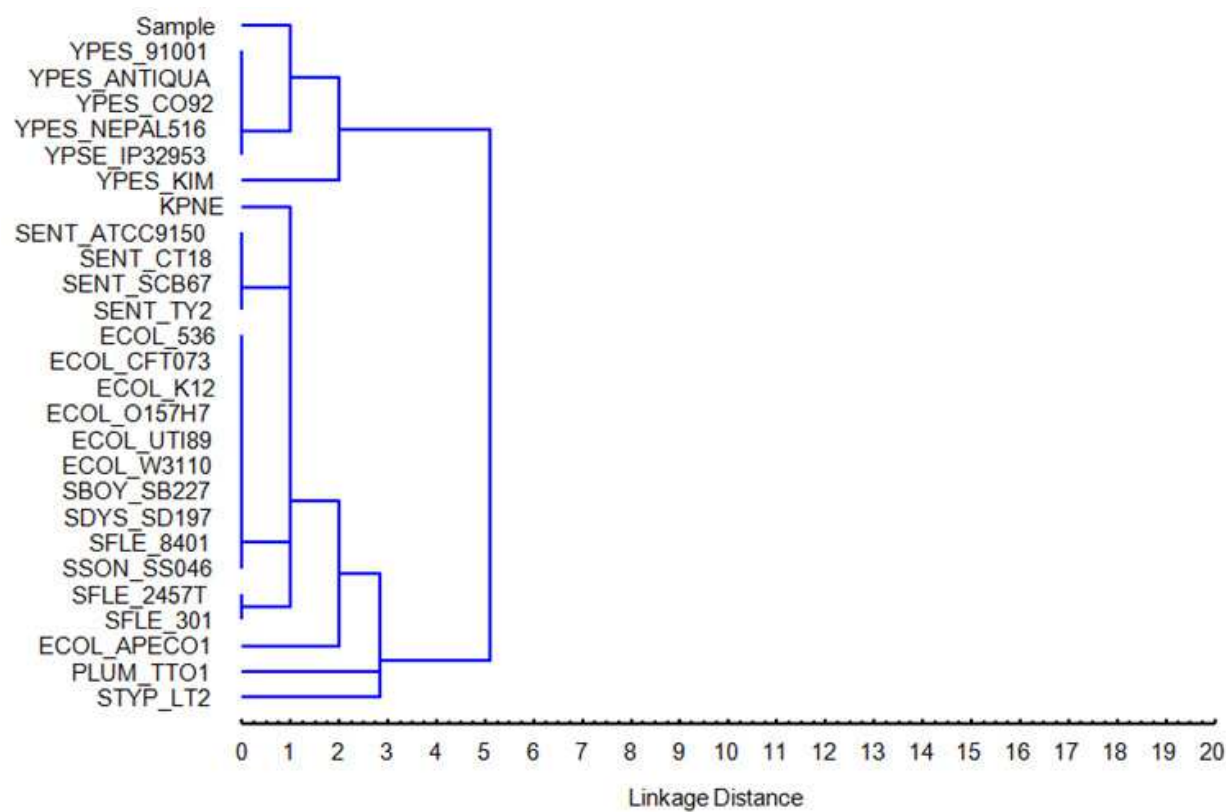
Fig. 5. Euclidean linkage similarity dendogram of the Nearest-neighbor classification of whole cell extract of *Yersinia pestis* CO92 strain.

Figure 6 shows the identification results for the OMP extracts of the *Y. pestis* CO92 sample. The dendrogram indicates an unambiguous, and correct, strain level identification with the *Y. pestis* CO92 strain in the proteome database. The experimental sample and Y. pestis entry of the *Y. pestis* CO92 strains are one linkage distance unit from the next nearest neighbor group consisting of the 91001/Antiqua/Nepal 516 strains. The set of unique proteins for virulent *Y. pestis* CO92 provides the presence of known biomarkers associated with virulence factors (Table 3). For example, virulence plasmids in *Y. pestis* such as pPCP1 that encodes for plasminogen activator protease precursor, pCD1 that encodes for low-calcium response protein, pMT1 that encodes for toxin protein and the structural gene for fraction 1 protein capsule (chaperonin protein) were found in the mass spectral analyses and are listed in Table 3. The chaperonin protein was present in higher abundance than that of the other protein biomarkers. The unique set of proteins had the closest match with *Y. pestis* strains compared to other similar bacteria in the database as seen in both dendrograms in Figures 5-6.

From analyses of both protein extracts, a comparison of the number of total, experimentally-determined number of proteins showed a difference between the two protein methods as applied to the *Y. pestis* sample. The whole cell protein and OMP approaches had 182 and 136, respectively, total identified proteins (Table 3). Upon removing the highly conserved, house-keeping, and energy transfer proteins from both strains, the number of strain-unique proteins (Table 3) for the whole cell protein and OMP approaches was four and thirteen, respectively. Even with a significant amount of unique proteins, the OMP differentiation capability did not provide a significant benefit (1.4 linkage units) with respect to the four

proteins from the whole cell approach (1 linkage unit) as detailed in the dendograms in Figures 5-6.

Whole Cell Extract		OMP Extract	
Accession Number	Protein Name	Accession Number	Protein Name
NP_993129	Hypothetical protein YP_1779	CAL19718	Cationic 19 kDa OMP
NP_995559	Murine toxin	NP_991899	Fraction 1 protein capsule (chaperonin GroEL)
NP_994104	Periplasmic chaperone	YP_070861	Membrane bound lytic murein transglycosylase C precursor
NP_991935	30S ribosomal protein S6	NP_993916	Aminotransferase
		NP_395168	Low-calcium response protein
		CAL18706	Secreted thiol: disulfide interchange protein DsbA
		CAL18984	Tellurium resistance protein
		CAL19717	Putative surface antigen
		CAL21872	Putative sigma 54 modulation protein
		NP_395233	Plasminogen activator protease precursor
		CAL19882	OMP porin C
		NP_395420	Murine toxin
		YP_02420	Probable formyl transferase
Total Identified Proteins	182		136

Table 3. Identified unique proteins lists detected in the whole cell and OMP extracts of *Y. Pestis* CO92 strain.
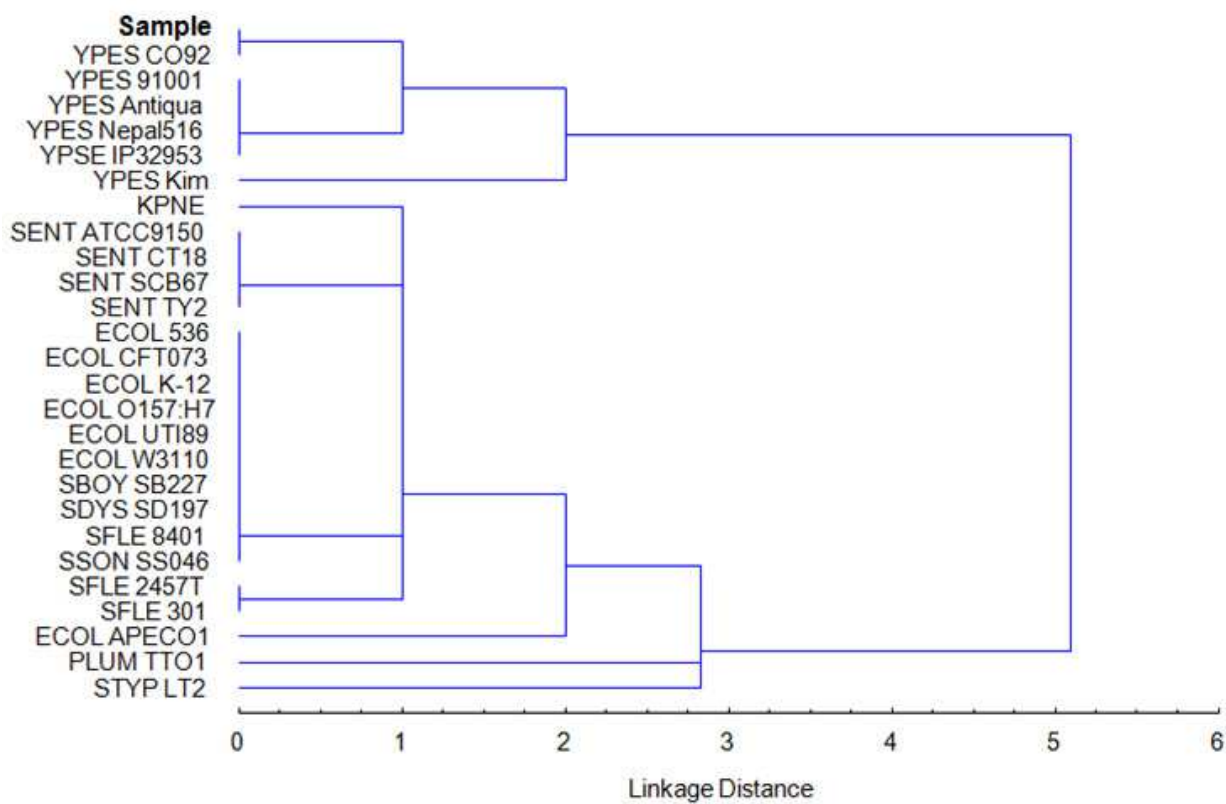
Fig. 6. Euclidean linkage similarity dendogram of the Nearest-neighbor classification of OMPs extract of *Yersinia pestis* CO92 strain.

6.4 Comparative proteomic differentiation between the whole cell and the OMP extracts for the *Y. pestis* A1122 strain

A comparison of the LC-Tandem MS and bioinformatics results of the proteins present in the whole cell and OMP extracts of the avirulent *Y. pestis* A1122 was performed. Figure 7 shows the nearest-neighbor similarity linkage analysis of the whole cell extract of the avirulent *Y. pestis* A1122 strain. A unique set of proteins for each extraction method had the closest match with *Y. pestis* strains compared to other similar Gram-negative bacteria in the database entries. In figure 7, the dendogram shows the similarity linkage for the whole cell protein extract from the *Y. pestis* A1122 in which the sample was identified to the pathogenic KIM, CO92 and Nepal 516 strains. Equidistant next nearest neighbors to this group are the 91001 and Antiqua strains. The linkage distance is minimal between these two groups of *Y. pestis* strains. On the basis of these results, the unique set of proteins (Table 4) from the experimental *Y. pestis* A1122 sample produced a closest similarity index to the CO92 and Nepal 516 virulent strains from whole cell protein extract preparations. A similar situation also was observed using whole cell protein extracts between the sample CO92 strain and the 91001/Antiqua/CO92/Nepal 516/IP32953 strains (Figure 5). As shown in table 4, there are three strain-unique proteins that were identified out of a total of 164 proteins from an analysis of the A1122 strain.
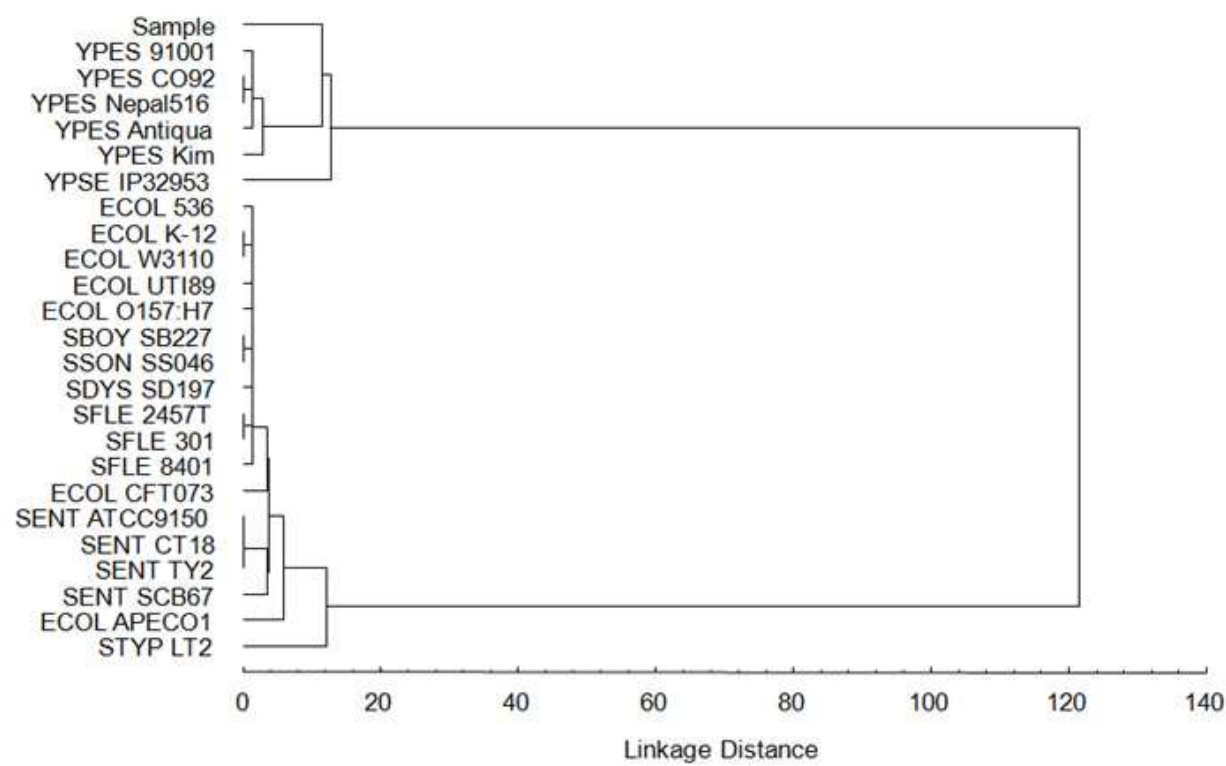
Fig. 7. Euclidean linkage similarity dendogram of the Nearest-neighbor classification of whole cell extract of *Yersinia pestis* A1122 strain.

On the other hand, the OMP analysis in Figure 8 shows that the sample was identified at the strain level as *Y. pestis* 91001. This finding is encouraging knowing that *Y. pestis* 91001 is the only avirulent strain in the proteome database which also includes several pathogenic *Y. pestis* strains. Because the avirulent *Y. pestis* A1122 strain has not been sequenced or is not publicly available, its absence from the database provided an indirect test of the robustness of the proteomics approach in the classification of a non-database bacterium to the database entries. It is worth mentioning that the constructed proteome data base consists of more than 1400 fully sequenced bacteria that had been translated into their complimentary protein expressions. All the samples studied were compared to all the proteomes in the constructed database and the top 20 closest near-neighbors were selected for further comparative proteomics analyses. This also provides confidence for identification at the species level (Figure 8). However, an equal similarity index is also shared with the Nepal 516 strain. The Antiqua strain is a very close nearest neighbor to the 91001 and Nepal 516 cluster of strains. The CO92 strain is observed to be relatively more removed from the 91001/Nepal 516 and Antiqua strains. On the basis of these results, the unique set of proteins for the experimental *Y. pestis* A1122 sample produced the same similarity index for the database *Y. pestis* 91001 and the Nepal 516 strains from the OMP extract preparation (Table 4). Figure 8 shows that there is a very small linkage distance between the groups of *Y. pestis* strains. Thus, the OMP analysis produces very similar classification results (very small linkage distances) for the six *Y. pestis* strains in the genome database. Table 4 lists the six unique proteins from a total of 94 proteins for the *Y. pestis* 91001 strain found in the OMP extract of the experimental A1122 strain. From

analyses of the whole cell protein extracts, a comparison of the total number of proteins produced 182 (Table 3) and 164 (Table 4) for *Y. pestis* CO92 and *Y. pestis* A1122, respectively. Upon removing the highly conserved, housekeeping, and energy transfer proteins from both strains, the number of strain-unique proteins was four for *Y. pestis* CO92 and three for *Y. pestis* A1122. From analyses of the OMP protein extracts, a comparison of the number of total, experimentally determined number of proteins showed a difference between *Y. pestis* CO92 and *Y. pestis* A1122. The CO92 strain had 136 total identified proteins compared to 94 for the A1122 strain. Upon removing the highly conserved, housekeeping, and energy transfer proteins from both strains, the number of strain-unique proteins for *Y. pestis* CO92 was 13 and that for *Y. pestis* A1122 was 6.
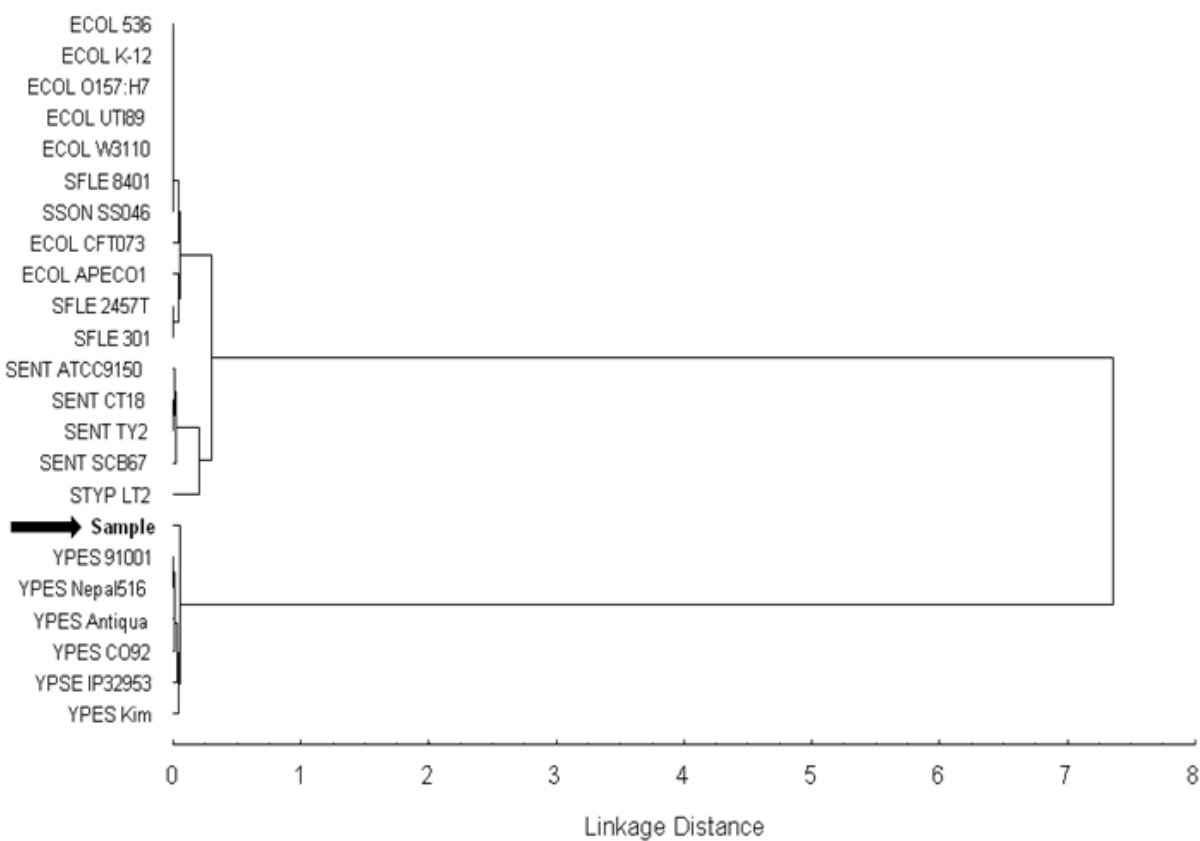

Fig. 8. Euclidean linkage similarity dendogram of the Nearest-neighbor classification of OMPs extract of *Yersinia pestis* A1122 strain.

Whole Cell Extract		OMP Extract	
Accession #	Protein Name	Accession #	Protein Name
NP_991849	Tellurium resistance protein	NP_991979	Transcription elongation factor NusA
NP_993230	arYPES_91001nyl-tRNA synthetase	NP_992082	Na(+)-translocating NADH-quinone reductase
NP_992224	Putative thioredoxin	NP_992120	Proline permease transport protein
		NP_993064	OMP porin
		NP_991484	Exported sulfate-binding protein
		NP_993650	OMP X
Total identified Proteins	164	94	

Table 4. Identified unique proteins lists detected in the whole cell and OMP extracts of *Y. Pestis* A1122 strain.

7. Conclusion

Comparative proteomics of tandem mass spectrometry data showed that the OMPs extract provided equal or better discrimination compared with the whole cell one with respect to the distance or similarity linkage with the next nearest neighbor(s). Also, the OMPs extracts of all studied strains showed correct database bacterial match with linkage similarity improved over the whole cell extract. The improved strain level differentiation using OMPs extract could be due to the possible ionization suppression experienced by whole cell that could shield the detection of important peptides that could be classified as unique biomarkers. However, whole cell lysates can be an appropriate option for the differentiation of Gram positive bacterial strains and the reported results herein support their potential application in bacterial species and potential strain differentiation. Also, Inclusion of more relevant bacteria such as *Francisella tularensis, Burkholderia*, and other Gram negative genera and species may provide a more comprehensive outlook on the importance of OMPs in comparison to the whole cell extract. These additions may also provide decision information as to the relative merit of applying OMP vs. whole cell protein extraction techniques in the analysis of an experimental bacterial sample for classification and diagnostic purposes.

Overall, Tandem MS-based proteomics and bioinformatics were shown to have utility in the comparative proteomics study for the differentiation of Gram-negative bacterial strains. Different numbers of distinguishing, unique proteins were obtained by the bioinformatics procedure between the whole cell and OMPs extracts. This resulted in different degrees of separation between the correctly determined database organism and the next nearest neighbor organism(s). Moreover, this approach relies on taxonomic correlation within the constructed proteome database and thus inferring an ID on sample organism not present in the genome database is possible. This capabilities is supported the fact that prokaryotic organism as they are arranged in hierarchal order their common proteins increase as we move from strain to phyla and vice versa. Such properties will allow the utilization of this

approach to infer taxonomic class based on the depth of available genomic sequencing information for such strains, i.e. species vs. genus vs. family vs. order, etc.

8. References

Aebersold, R. (2003).A mass spectrometric journey into protein and proteome research, *J. Am. Soc. Mass Spectrom.* 14, 685-695.

Anhalt, J.P. & Fenselau, C. (1975). Identification of Bacteria using mass spectrometry, *Anal. Chem.* 47,219-225.

Baglioni, P.; Bini, L.; Liberatori, S.; Pallini, V.& Marri, L. (2003). Analysis of *E. coli* W3110 expressing heterologous sigma factor reported. *Proteomics*, 3, 1060-1065.

Ben-Gurion, R.; Shafferman, A. (1981). Essential virulence determinants of different *Yersinia* species are carried on a common plasmid. *Plasmid* , 5, 183-187.

Bevilacqua, V. L. H.; Nilles, J. M.; Rice, J. S.; Connell, T. R. ; Schenning, A. M.; Reilly, L. M.& Durst, H. D. (2010). Ricin Activity Assay by Direct Analysis in Real Time Mass Spectrometry Detection of Adenine Release. *Anal. Chem.*, 82(3): 798–800.

Castanha, E. R.; Fox, A.& Fox, K. F. (2006). Rapid discrimination of *Bacillus anthracis* from other members of the *B. cereus* group by mass and sequence of "intact" small acid soluble proteins (SASPs) using mass spectrometry. *J. Microbiol. Methods*, 67, 230-240

Chen, W.; Laidig, K. E.; Park, Y.; Park, K.; Yates III, J. R.; Lamont, R. J.& Hackett, M. (2001). Searching the *Porphyromonas gingivalis* genome with peptide fragmentation mass spectra. *Analyst*, 126, 52-57.

Cornelis, G. R. (2002). *Yersinia* type III secretion: send in the effectors. *J. Cell Biol.*, 158, 401-408.

Costas; M. (1990). Numerical-Analysis of sodium dodecyl sulfate-polyacrylamide gel-electrophoretic protein-patterns for the classification, identification and typing of medically important bacteria. *Electrophoresis* 11: 382–391.

Craig, R.; & Beavis, R. (2004). TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*, 20:1466-1467.

Demirev, P. A. & Fenselau, C. (2008). Mass Spectrometry for Rapid Characterization of Microorganisms. *Annu. Rev. Anal. Chem.*, 1, 71-93.

Demirev, P. A. & Fenselau, C. (2008). Mass spectrometry in biodefense. *J. Mass Spectrom.*, 43, 1441-1457

Demirev, P. A.; Feldman, A. B.& Lin, J. S. (2005). Chemical and biological weapons: current concepts for future defenses. *Johns Hopkins APL Tech. Digest*, 26, 321-333.

Dworzanski, J.P.; Deshpande, S.V.; Chen, R.; Jabbour, R.; Snyder, A.P.; Wick, C.H. & L. Li. (2006). Mass Spectrometry-Based Proteomics Combined with Bioinformatic Tools for Bacterial Classification, *Journal of Proteome Research*, 5, 76-87.

Dworzanski, J. P.& Snyder, A. P. (2005). Classification and identification of bacteria using mass spectrometry-based proteomics. *Expert Review Proteomics*, 2, 863-878.

Dworzanski, J. P.; Snyder, A. P.; Chen, R.; Zhang, H.; Wishart, D. & Li, L. (2004). Identification of Bacteria Using Tandem Mass Spectrometry Combined with a Proteome Database and Statistical Scoring, *Anal. Chem.*, 76, 2355-2366.

Ecker, D. F.; Sampath, R.; Blyn, L. B.; Eshoo, M. W.; Ivy, C.; Ecker, J. A.; Libby, B.; Samant, V.; Sannes-Lowery, K. A.; Melton, R. E.; Russell, K. Freed, N.; Barrozo, C.; Wu, J.; Rudnick, K.; Desai, A.; Moradi, E.; Knize, D. J.; Robbins, D. W.; Hannis, J. C.; Harrell, P. M.; Massire, C.; Hall, T. A.; Jiang, Y.; Ranken, R.; Drader, J. J.; White, N.; McNeil, J. A.; Crooke, S. T. & Hofstadler, S. A. (2005). Rapid identification and strain-typing of respiratory pathogens for epidemic surveillance. *Proc. Natl. Acad. Sci.*, 102, 8012-8017.

Eng, J. K.; McCormack, A. L. & Yates, J. R. (1994). An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *J. Am. Soc. Mass. Spectrom.* 5:976-989.

Fagerquist, C. K.; Garbus, B. R. ; Miller, W. G. ; Williams, K. E.; Yee, E.; Bates, A. H.; Boyle, S.; Harden, L. A.; Cooley, M. B.& Mandrell, R. E. (2010). Rapid Identification of Protein Biomarkers of *Escherichia coli* O157:H7 by Matrix-Assisted Laser Desorption Ionization-Time-of-Flight–Time-of-Flight Mass Spectrometry and Top-Down Proteomics. *Anal. Chem.*, *82*, 2717-2725.

Ferber, D. M.& Brubaker, R. R. (1981). Plasmids in *Yersinia pestis*. *Infect. Immun.*, *31*, 839-841.

Filippov, A. A.; Solodovnikov, N. S.; Kookleva, L. M.; Protsenko, O. A. (1990). Plasmid content in Yersinia pestis of different origin. *FEMS Microbiol. Lett.*, *67*, 45-48.

Fox, A. (2006). Mass Spectrometry for Species or Strain Identification after Culture or without Culture: Past, Present, and Future. *J. Clin. Microbiol.*, *44*, 2677-2680.

Fox, A.; Anderson, M.; Dunn, J.; Guenther, B.; Parks, L.; Pinnick, R.; Reed, C.; Rowe, J.; Luftig, R.; Long, G. W.; Lontz, R. Marchin, G. L.; McManus, A. T.; Setlow, P.; Siedow, J.; Sievers, A. J.; Tchikindas, M. L.& Tove, S. (2002) Report of the "Bioterrorism Workshop"—Duke University Thomas Center on 2–4 April 2002 organized by US Army Research Office. *J. Microbiol. Methods*, *51*, 247-254.

Friess, R. (2008). The New York Times; August 5, 2008; http://www.nytimes.com/2008/08/05/us/05ricin.html?scp=1&sq=ricin&st=nyt. Accessed February 12, 2010

Garmendia, J.; Frankel, G. & Crepin, V. F. (2005). Enteropathogenic and Enterohemorrhagic *Escherichia coli* Infections: Translocation, Translocation, Translocation. *Infect. Immun.*, *73*, 2573-2585.

Hillenkamp, F.& Karas, M. (2000).Matrix-assisted laser desorption/ionization, an experience, *Int. J. Mass Spectrom.*, 200, 71-77.

Ho, Y.-P. & Hsu, P.-H. (2002).Investigating the effects of protein patterns on microorganism identification by high-performance liquid chromatography-mass spectrometry and protein database searches. *J. Chromatogr. A*, *976*, 103-111

Hofstadler, S. A.; Sampath, R.; Blyn, L. B.; Eshoo, M. W.; Hall, T. A.; Jiang, Y.; Drader, J. J.; Hannis, J. C.; Sannes-Lowery, K. A., Cummins, L. L.; Libby, B.; Walcott, D. J.; Schink, A.; Massire, C.; Ranken, R.; Gutierrez, J.; Manalili, S.; Ivy, C.; Melton, R.; Levene, H.; Barrett-Wilt, G.; Li. F.; Zapp, V.; White, N.; Samant, V.; McNeil, J. A.; Knize, D.; Robbins, D.; Rudnick, K.; Desai, A.; Moradi, E. & Ecker, D. J. (2005). TIGER: the universal biosensor, *Intl. J. Mass Spectrom.*, 242, 23-41.

http://www.ncbi.nlm.gov/PMGifs/genomics/micr/html

http://wit.integratedgenomics.com/GOLD

Ide, T., Laarmann, S.; Greune, L.; Schillers, H.; Oberleithner, H.; Schmidt, M. A. (2001). Characterization of translocation pores inserted into plasma membranes by type III-secreted Esp proteins of enteropathogenic *Escherichia coli*. *Cell. Microbiol*. 3, 669-679

Jabbour, R.; Dworzanski, J.P.; Deshpande, S.V.; Snyder, A.P. & Wick, C.H. (2005).Effect of gas phase fractionation of peptide ions on bacterial identification using mass spectrometry-based proteomics approach. *Proceedings of the 53rd ASMS Conference on Mass Spectrometry and Allied Topics*, TX, USA, TP31

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

Jerse, A.E.; Yu, J.; Tall, B.D.; Kaper, J.B.; (1990). A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. *Proc Natl Acad Sci USA.*;87:7839–43.

Kaper, J.B.; Nataro, J.P.; Mobley, H.L.T. (2004). Pathogenic *Escherichia coli*. *Nat Rev Microbiol*; 2,123–138.

Keller, A.; Nesvizhskii, A. I.; Kolker, I. & Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS-MS and database search. *Anal. Chem.* 74:5383-5392

Koebnik, R.; Locher, K.P.; Van Gelder, P. (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol.,*37:239–253

Krishnamurthy, T.; Rajamani, U.; Ross, P. L.; Jabbour, R.; Nair, H.; Eng, J.; Yates, J.; Douglas, M. T. D.; Stahl, C. & Lee, T. D. (2000). Mass spectral investigation of microorganisms, *Toxin Reviews,*19:1, 95-117

Kuwana, R.; Kasahara, Y.; Fujibayashi, M.; Takamatsu, H.; Ogasawara, N.& Watabe, K. (2002). Proteomics characterization of novel spore proteins of *Bacillus subtilis*. *Microbiology*, *148*, 3971-3982.

Lambert, J.-P.; Ethier, M.; Smith, J. C.& Figys, D. (2005) Multiplexed Proteomic Reactor for the Processing of Proteomic Samples. *Anal. Chem.*, *77*, 3771-3788.

Lamontagne, J.; Butler, H.; Chaves-Olarte, E.; Hunter, J.; Schirm, M.; Paquet, C.; Tian, M.; Kearney, P.; Hamaidi, L.; Chelsky, D.; Moriyon, M.; Moreno, E.& Paramithiotis, E. (2007). Extensive cell envelope modulation is associated with virulence in Brucella abortus. *J. Prot. Res.*, *6*, 1519-1529.

Loferer-Krobacher, M.; Klima, J. & Psenner R. (1998). Determination of Bacterial Cell Dry Mass by Transmission Electron Microscopy and Densitometric Image Analysis, *Applied Environ. Microbiol.*, 64, 688–694.

Lipton, M.S., Pasa-Tolic, L.; Anderson, G. A.; Anderson, D. J.; Auberry, D. L. ; Battista, J. R. ; Daly, M. J.; Fredrickson, J.; Hixson, K. K.; Ksotandarithes, H.; Masselon, C.; Markillie, L. M.; Moore, R. J.; Romine, M. F.; Shen, Y.; Stritmatter, E.; Tolic, N.; Udseth, H. R.; Venkaeswaran, A.; Wong, K. K.; Zhao, R.& Smith, R. D. (2002). Global analysis of the *Deinococcus radiodurans* proteome by using accurate mass tags. *Proc. Natl. Acad. Sci. U.S.A.* 99;11049-11054.

Ma, J.; Liu, J.; Sun, L.; Gao, L.; Liang, Z.; Zhang, L.& Zhang, Y. (2009). Online Integration of Multiple Sample Pretreatment Steps Involving Denaturation, Reduction, and Digestion with Microflow Reversed-Phase Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry for High-Throughput Proteome Profiling *Anal. Chem.*, *81*, 6534-6540.

McDaniel, T. K.& Kaper, J. B. (1997). A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on E. coli K-12. *Mol. Microbiol.*, *23*, 399-407.

McDaniel, T. K.; Jarvis, K. G.; Donnenberg, M. S.; Kaper, J. B. (1995). A Genetic Locus of Enterocyte Effacement Conserved among Diverse Enterobacterial Pathogens. *Proc. Natl. Acad. Sci. USA*, *92*, 1664-1668.

Nagele, E.; Vollmer, M.; Horth, P. (2003). Two-dimensional nano-liquid chromatography–mass spectrometry system for applications in proteomics. *J. Chromatogr. A*, *1009*, 197-205

National Research Council. (2005). Sensor Systems for Biological Agent Attacks. Washington, DC: Natl. Acad. Press, ISBN-10: 0-309-09576-X.

Norbeck, A.D.; Callister, S. J. ; Monroe, M. E.; Jaitly, N.; Elias, D. A.; Lipton, M. S.& Smith, R. D. (2006). Proteomic approaches to bacterial differentiation. *J. Microbiol. Methods* 67,473-486.

Pennington, S. R.; Wilkins, M. R.; Hochstrasser, D. F. & Dunn, M. J. (1997) Proteome analysis: from protein characterization to biological function. *Trends Cell Biol.*, 7, 168-173.

Pribil, P.; Fenselau, C. (2005). Characterization of Enterobacteria Using MALDI-TOF Mass Spectrometry. *Anal. Chem., 77,* 6092-6095.

Ramamurthi, K. S.& Schneewind, O. (2002). Type III protein secretion in *Yersinia* species. *Annu. Rev. Cell Dev. Biol., 18,* 107-133.

Rotz, L. D.; Khan, A. S.; Lillibridge, S. R.; Ostroff, S. M.& Hughes, J. M. (2002). Bioterrorism Preparedness for the Public Health and Medical Communities. *Emerging Infect. Dis., 8,* 225-230.

Sampath, R.; Hall, T. A.; Massire, C.; Li, F.; Blyn, L. B.; Eshoo, M. W.; Hofstadler, S. A. & Ecker, D. J. (2007). Rapid Identification of Emerging Infectious Agents Using PCR and Electrospray Ionization Mass Spectrometry. *Annals New York Acad. Sci*., 1102, 109-120.

Titball, R. W.; Hill, J.; Lawton, D. G. & Brown, K. A. (2003). *Yersinia pestis* and plague *Biochem. Soc. Trans., 31,* 104-107.

VerBerkmoes, N.C.; Hervey, W. J.; Shah, M.; Land, M.; Hauser, L.; Larimer, F. W.; van Berkel, G. J. & Goeringer, D. E. (2005). Evaluation of "shotgun" proteomics for identification of biological threat agents in complex environmental matrices: experimental simulations. *Anal. Chem.,* 77:923-932.

Wachter, C.; Beinke, C.; Mattes, M. & Schmidt, M. A. (1999). Insertion of EspD into epithelial target cell membranes by infecting enteropathogenic Escherichia coli. *Mol. Microbiol., 31,* 1695-1702.

Warscheid, B, and C. Fenselau. (2003). Characterization of *Bacillus* spore species and their mixtures using post source decay with a curved-field reflectron. *Anal. Chem.* 75:5618-5627.

Warscheid, B.; and Fenselau, C. (2004). A targeted proteomics approach to the rapid identification of bacterial cell mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *Proteomics* 4:2877-2892.

Washburn, M. P.; Wolters, D. & J. R. Yates III. (2001). Large-scale analysis of yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19:242-247.

Wilkins, C. & Lay, J. (Eds.). (2006). Identification of Microorganisms by Mass Spectrometry. John Wiley & Sons: Hoboken; ISBN: 978-0-471-65442-1.

Williams, T. L.; Leopold, P.& Musser, S. (2002). Automated post processing of electrospray LC/MS data for profiling protein expression in bacteria *Anal. Chem., 74,* 5807-5813.

Wolters, D. A., Washburn, M. P. & J. R. Yates III. (2001). An automated multidimensional protein identification technology for shotgun proteomics. Anal. Chem. 73:5683-5690.

Wu, C. C.; MacCoss, M. J.; Howell, K. E.&Yates III, J. R. (2003). A method for the comprehensive proteomic analysis of membrane proteins. *Nat. Biotechnol., 21,* 532-538.

Xiang, F.; Anderson, G.; Veenstra, T.; Lipton, M.& Smith, R. (2000). Characterization of microorganisms and biomarker development from global ESI-MS-MS analyses of cell lysates. *Anal. Chem.,* 72:2475-2481.

Yamada, M.; Asaoka, S.; Saier, Jr., M. H. & Yamada, Y. (1993). Characterization of the gcd gene from Escherichia coli K-12 W3110 and regulation of its expression. *J. Bacteriol.,* 175, 568-571.

Yates, J. R. (1998). Mass Spectrometry and the Age of the Proteome. *J. Mass Spectrom., 33,* 1-19.

Zheng, S.; Schneider, K. A.; Barder, T. J.& Lubman, D. M. (2003). Two-dimensional liquid chromatography protein expression mapping for differential proteomic analysis of normal and O157:H7 Escherichia coli. *BioTechniques, 35,* 1202-1212.
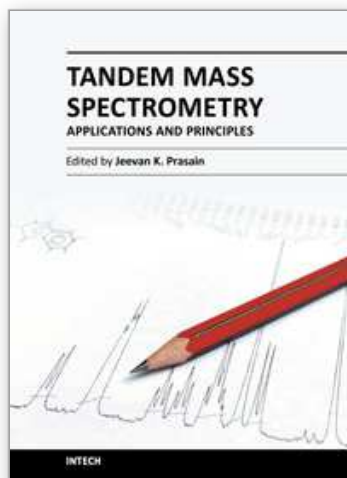
Tandem Mass Spectrometry - Applications and Principles

Edited by Dr Jeevan Prasain

Tandem Mass Spectrometry - Applications and Principles presents comprehensive coverage of theory, instrumentation and major applications of tandem mass spectrometry. The areas covered range from the analysis of drug metabolites, proteins and complex lipids to clinical diagnosis. This book serves multiple groups of audiences; professional (academic and industry), graduate students and general readers interested in the use of modern mass spectrometry in solving critical questions of chemical and biological sciences.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Rabih E. Jabbour, Mary M. Wade, Samir V. Deshpande, Michael F. Stanford, Alan W. Zulich and A. Peter Snyder (2012). Comparative Proteomics of Tandem Mass Spectrometry Analyses for Bacterial Strains Identification and Differentiation, Tandem Mass Spectrometry - Applications and Principles, Dr Jeevan Prasain (Ed.), ISBN: 978-953-51-0141-3, InTech, Available from: http://www.intechopen.com/books/tandem-mass-spectrometry-applications-and-principles/comparative-proteomic-for-bacterial-differentiation-using-tandem-mass-spectrometry

INTECH
open science | open minds