Tips & Tools

Teaching Microbial Identification with Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) and Bioinformatics Tools †

Wenfa Ng

Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117576

INTRODUCTION

Microbial identification, together with studies investigating the microorganism's physiology, genetics, metabolism and ecological functions, comprise the intellectual pillars of microbiology (12). Currently, polymerase chain reaction (PCR) amplification and sequencing of 16S rRNA genes (20) is the gold-standard approach for identifying unknown microorganisms in clinical and environmental samples. Though PCR lends itself to the relatively fast culture-independent identification of microorganisms, the extensive and complex sample preparation required, as well as expensive reagents, has motivated the search for PCR-independent methods (7). One such method is matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), which, although not culture-independent, provides a rapid and relatively low-cost (5) method for identifying bacteria, yeast, and fungi, often to the species and, on occasion, sub-species level, with little or no sample preparation (1, 17). These characteristics enable its clinical use for screening and identifying known pathogenic microorganisms for disease control and treatment (17). Given its increasing use in clinical diagnostics (3, 4) and potential for application in the food industry (13), the technique's conspicuous absence in the undergraduate curriculum presents a curriculum gap that can be filled via exercises utilizing this commercialized technique.

Pedagogical research has highlighted the beneficial effects of incorporating activities—for example, simple demonstrations, experiments, and field trips—into classes to: (i) create interest, and thus, motivate students to learn the subject matter, and (ii) help students make important connections between textbook learning and real-world phenomena (2, 9–12, 22). This article describes a laboratory exercise (suitable for complementing a variety of life sciences, analytical chemistry, and bioinformatics courses) that involve students in collecting water samples from a drain or small stream on campus, cultivating, and subsequently identifying the microorganisms in the water using a combined MALDI-TOF MS and bioinformatics approach. By blending modern mass spectrometry instrumentation and bioinformatics concepts for phylogenetic analysis with cell culture techniques, this exercise provides students with the opportunity to learn leading-edge concepts and gain hands-on experience of instrumentation that they will likely encounter in their future bioscience careers or in graduate and medical school. More important, in contrast to laboratory experiments designed to verify known observations or theories, the campus "field trip" to collect real water samples can help to open the students' minds to the unknown as well as kindle their curiosity and spirit of inquiry—skill sets critical for leading an intellectually fulfilling life in a knowledge-based economy. Finally, with slight modifications to sample collection and inoculation procedures, this activity is broadly applicable to performing a census of microbes in diverse contexts such as on the surfaces of vegetables or a person's skin.

PROCEDURE

The laboratory exercise comprises four main segments: (i) sample collection, (ii) cultivation, (iii) MALDI-TOF MS, and (iv) bioinformatics analysis (Fig. 1). Firstly, 100 ml pre-sterilized polypropylene bottles (Nalgene, USA) were used to collect environmental water samples that, in turn, served as inoculum for pre-prepared R2A agar plates (Difco, USA; composition in g/L: yeast extract, 0.5; proteose peptone No. 3, 0.5; casamino acids, 0.5; D-glucose, 0.5; soluble starch, 0.5; sodium pyruvate, 0.3; dipotassium phosphate, 0.3; magnesium sulfate, 0.05; agar, 15.0)—the gold-standard for cultivating microorganisms from oligotrophic environments (14). The water's pH, temperature, and conductivity (a measure of salinity) should also be measured with a portable pH/conductivity meter if possible. Using the spread plate technique under aseptic conditions, 0.1 or 1 ml of the water sample is spread onto R2A agar (3 technical replicates) with a sterile plastic disposable spreader. As the water sample may contain vegetative cells and spores from a variety of

Corresponding author. Mailing address: Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117576. E-mail: ngwenfa@alumni.nus.edu.sg.
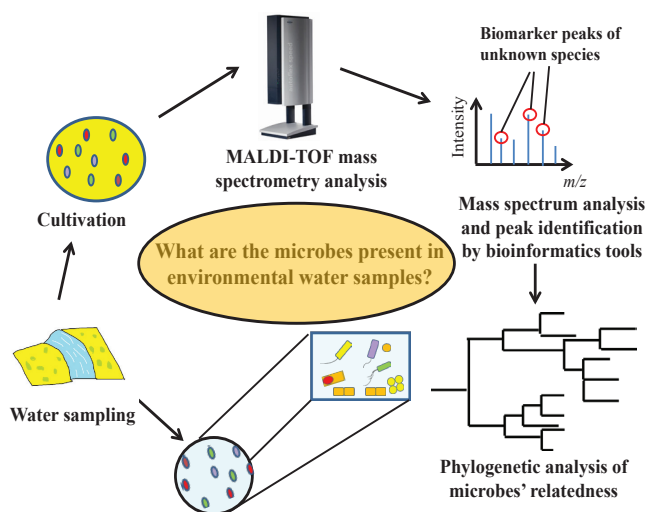
FIGURE 1. Overview of the various components—and key learning objectives—of the laboratory exercise. Combining cell culture, mass spectrometry, and bioinformatics analysis of biomarker peaks from cultivated microorganisms, students engage in an inquiry-based exercise using the scientific method to tackle a research question with unknown answers.
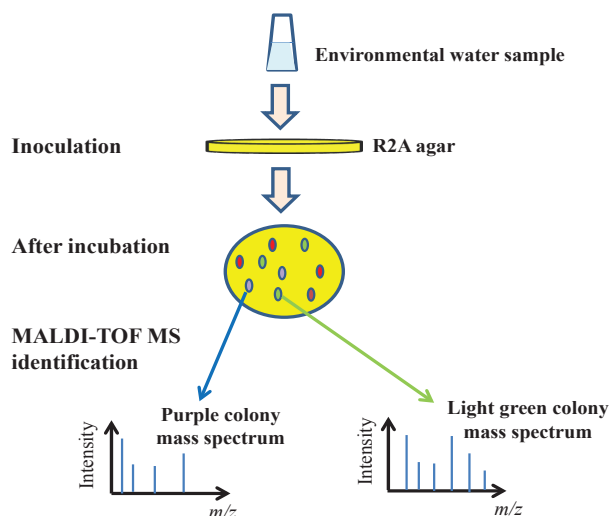

FIGURE 2. Concept underlying MALDI-TOF MS-based identification of microorganisms, in particular, differences in *m/z* ratios and relative intensities of the mass peaks afford the ability to distinguish between species. Note the need for a culture step prior to mass spectrometry analysis, a key difference from the culture-independent approach of PCR amplification followed by 16S rRNA sequencing.

microorganisms, the inoculation step should be performed in a Class II biological safety cabinet. The inoculated plates are subsequently incubated in an inverted position at a temperature mimicking that of the sample collection site. Depending on the incubation temperature, nutrient conditions of the water sample, and the microbial species present, a couple of days are generally needed for cultivation (especially for slow-growing microorganisms), but students should check the agar plates daily to observe the microorganisms' relative growth rates—information important to hypothesizing metabolic and functional relationships between them.

After a few days of incubation, colonies of at least 0.4 cm in diameter can be identified by MALDI-TOF MS (Fig. 2) (see Appendix 1), which has a detection limit of $\sim 10^3$ cells/ml (6). Briefly, a 10 µl plastic inoculation loop is used to directly smear a colony from the agar onto a 384-well stainless steel target plate. A small volume (1–2 µl) of MALDI matrix (either α-Cyano-4-hydroxycinnamic acid (CHCA) or 2,5-dihydroxybenzoic acid (DHB) dissolved in an acetonitrile (50%) and trifluoroacetic acid (2.5%) buffer) is subsequently overlaid onto the sample, and the mixture allowed to air-dry for 15 minutes at room temperature before analysis by a MALDI-TOF MS operated under linear positive ion mode with a nitrogen UV laser (337 nm). Both CHCA and DHB matrixes offer good ionization of biomolecules in the mass range from 2,000 to 20,000 Daltons (Da). Prior to analysis, the instrument should be calibrated with an external standard, for example, cytochrome c (~12,000 Da). For mass spectrum acquisition, 50 laser pulses are directed at each spot of the sample mixture; however, due to the presence of "sweet spots" in the mixture that typically produce better mass spectra,

interrogation of a number of spots is needed to identify an ideal sample location for generating high quality mass spectra (16). Although the analysis process can be automated, research reveals that manual spectra acquisition would, in general, generate better spectra (15) and allow students to experience all aspects of instrument operation—from sample preparation to optimization of analysis conditions, and finally, spectra acquisition.

After collection of mass spectrum data, students identify the microorganisms using the proteome database search approach (8, 21). Briefly, the ribosomal protein sequences—translated from those of the corresponding genes—of various microbes whose sequenced genomes are deposited in the Microbial Genome Database for Comparative Analysis (MBGD) (http://mbgd.genome.ad.jp/) (19) are queried, and their masses—and thus, *m/z* ratios—obtained from the protein's UniProt page or calculated based on that of the constituent amino acids (assuming no fragmentation and an overall molecular charge of +1). Next, the *m/z* ratios are used to search for their corresponding mass peaks in the extracted peak lists of various mass spectra for identification purposes (Fig. 3). Additionally, construction of a phylogenetic tree encompassing all identified microorganisms could be undertaken via the Molecular Evolutionary Genetics Analysis (MEGA) software tool (http://www.megasoftware.net/) (18) to provide a visual map of the sample's microbial diversity and the microorganisms' relatedness. Furthermore, hypotheses postulating the microorganisms' ecological functions and interactions can be formulated based on their growth characteristics (for example, closely associated colonies on agar) and their habitat's environmental conditions.

Species: *Haemophilus influenzae* Rd KW20
Description: 30S ribosomal protein S16
Protein name: RpsP

Nucleotide sequence → Amino acid sequence → Mass of ribosomal protein corresponds to its *m/z* by assuming +1 charge → Search for corresponding mass peak in MALDI-TOF mass spectrum

atggtaaccattcgtttatctc
gtggcggagctaaaaaacg
cccattttatcaaatcgta
gtcgctgacagtcgttcacct
cgtgacggtcgcttcattgaa
cgtgtaggtttcttcaat
ccaattgcacaaggtaatgc
agaacgtcttcgcattaatctt
gagcgtgtaaatcactgg
gttgctcaaggtgcttcacttt
ctgatcgtgttgcgagcttgg
taaaagaagcacaaaaa
gcagcataa

MVTIRLSRG
GAKKRPFYQ
IVVADSRSPR
DGRFIERVGF
FNPIAQGNA
ERLRINLERV
NHW
VAQGASLSD
RVASLVKEA
QKAA

m/z = 9179 Da

Molecular weight can be obtained via accessing the protein's UniProt page, or alternatively, it can be calculated by summing the mass of individual amino acid residues constituting the protein

Peak identification

Intensity

m/z

Note that a single annotated peak is not sufficient for microbial identification; given the uncertainty in mass resolution and protein sequence information

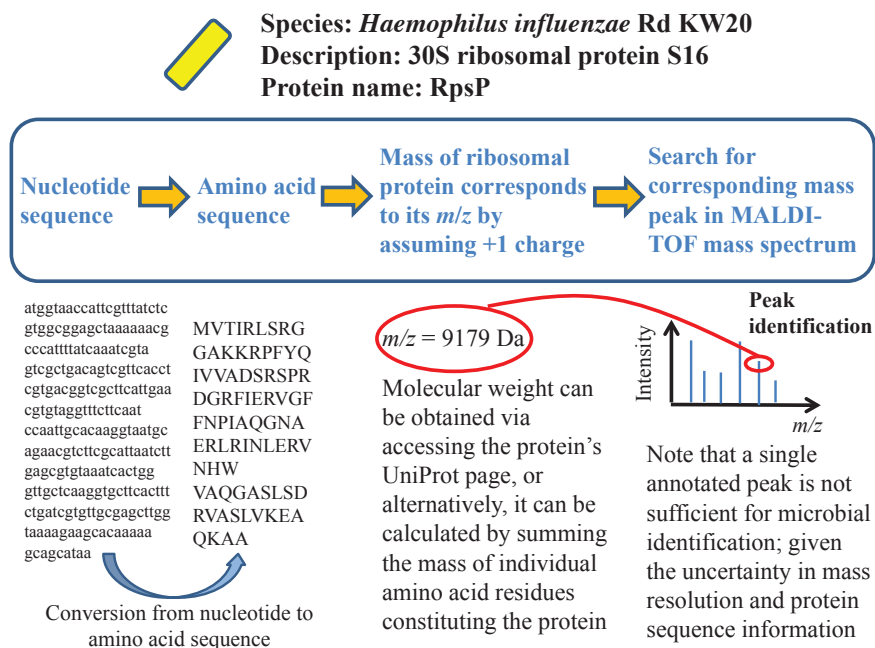Conversion from nucleotide to amino acid sequence

FIGURE 3. Steps important to the identification of mass peaks—using the proteome database search approach—from MALDI-TOF mass spectrum of individual colony. Briefly, MBGD stores both the nucleotide and amino acid sequence of each ribosomal gene—where accessing the gene's UniProt page yields the ribosomal protein's molecular weight (and also *m/z* ratio; assuming +1 molecular ion charge), which, in turn, can be used to search the extracted peaks list of each mass spectrum.

CONCLUSION

A curriculum is a living entity: specifically, it evolves and responds to changes in technology, pedagogy, and societal needs. Centrality of microbial identification in microbiology, emergence of MALDI-TOF MS as a rapid tool for identifying microorganisms, as well as the advent of bioinformatics algorithms capable of querying huge datasets, call for pedagogical innovations to impart, in each student, the ability to isolate and cultivate microorganisms from samples and competently use modern mass spectrometry instruments and bioinformatics tools to ask, search, and answer questions concerning species' identities. This communication describes an inquiry-based laboratory exercise that integrates the learning of theoretical and operational aspects of MALDI-TOF MS and bioinformatics tools through practical hands-on experimentation addressing a research question with unknown answers: what are the microorganisms present in an environmental water sample? From water sample collection, inoculation and cultivation, daily observations of growth behavior, MALDI-TOF MS analysis, and finally, interpretation of mass spectra and identification of microorganisms aided by bioinformatics tools, students can learn experimental and computational techniques in addition to practicing the scientific method in answering a research question. The most important learning outcome, however, is imbuing, in the students, a sense of wonder in the natural world and igniting their inquiring minds.

SUPPLEMENTAL MATERIALS

Appendix 1: Theoretical overview and practical tips for exercise implementation

ACKNOWLEDGMENTS

REFERENCES

1. **Anhalt, J. P., and C. Fenselau.** 1975. Identification of bacteria using mass spectrometry. Anal. Chem. **47:**219–225.
2. **Bascom-Slack, C. A., A. E. Arnold, and S. A. Strobel.** 2012. Student-directed discovery of the plant microbiome and its products. Science **338:**485–486.
3. **Bessède, E., et al.** 2011. Matrix-assisted laser-desorption/ionization biotyper: experience in the routine of a university hospital. Clin. Microbiol. Infect. **17:**533–538.
4. **Bizzini, A., C. Durussel, J. Bille, G. Greub, and G. Prod'hom.** 2010. Performance of matrix-assisted laser desorption ionization-time of flight mass spectrometry for

identification of bacterial strains routinely isolated in a clinical microbiology laboratory. J. Clin. Microbiol. **48:**1549–1554.

5. **Cherkaoui, A., et al.** 2010. Comparison of two matrix-assisted laser desorption ionization-time of flight mass spectrometry methods with conventional phenotypic identification for routine identification of bacteria to the species level. J. Clin. Microbiol. **48:**1169–1175.

6. **Croxatto, A., G. Prod'hom, and G. Greub.** 2012. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. FEMS Microbiol. Rev. **36:**380–407.

7. **DeAngelis, K. M., et al.** 2011. PCR amplification-independent methods for detection of microbial communities by the high-density microarray PhyloChip. Appl. Environ. Microbiol. **77:**6313–6322.

8. **Demirev, P. A., Y.-P. Ho, V. Ryzhov, and C. Fenselau.** 1999. Microorganism identification by mass spectrometry and protein database searches. Anal. Chem. **71:**2732–2738.

9. **Forest, K., and S. Rayne.** 2009. Thinking outside the classroom: integrating field trips into a first-year undergraduate chemistry curriculum. J. Chem. Educ. **86:**1290–1294.

10. **Gammon, K.** 2011. Room to grow—and learn. HHMI Bulletin **24:**38–39.

11. **Korfiatis, K. J., and S. D. Tunnicliffe.** 2012. The living world in the curriculum: ecology, an essential part of biology learning. J. Biol. Educ. **46:**125–127.

12. **Merkel, S.** 2012. The development of curricular guidelines for introductory microbiology that focus on understanding. J. Microbiol. Biol. Educ. **13:**32–38.

13. **Nicolaou, N., Y. Xu, and R. Goodacre.** 2012. Detection and quantification of bacterial spoilage in milk and pork meat using MALDI-TOF-MS and multivariate analysis. Anal. Chem. **84:**5951–5958.

14. **Reasoner, D. J., and E. E. Geldreich.** 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl. Environ. Microbiol. **49:**1–7.

15. **Schumaker, S., C. M. Borror, and T. R. Sandrin.** 2012. Automating data acquisition affects mass spectrum quality and reproducibility during bacterial profiling using an intact cell sample preparation method with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. **26:**243–253.

16. **Šedo, O., A. Voráč, and Z. Zdráhal.** 2011. Optimization of mass spectral features in MALDI-TOF MS profiling of *Acinetobacter* species. Syst. Appl. Microbiol. **34:**30–34.

17. **Seng, P., et al.** 2009. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin. Infect. Dis. **49:**543–551.

18. **Tamura, K., et al.** 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28:**2731–2739.

19. **Uchiyama, I.** 2003. MBGD: microbial genome database for comparative analysis. Nucleic Acids Res. **31:**58–62.

20. **Woese, C. R., and G. E. Fox.** 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. U. S. A. **74:**5088–5090.

21. **Wynne, C., C. Fenselau, P. A. Demirev, and N. Edwards.** 2009. Top-down identification of protein biomarkers in bacteria with unsequenced genomes. Anal. Chem. **81:**9633–9642.

22. **Zhang, X.** 2011. Teaching molecular phylogenetics through investigating a real-world phylogenetic problem. J. Biol. Educ. **46:**103–109.