JCM
Journals.ASM.org

Fusobacterium nucleatum Subspecies Identification by Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry

Shuping Nie,[a,b] Baoyu Tian,[c] Xiaowei Wang,[c] David H. Pincus,[f] Martin Welker,[g] Kathleen Gilhuley,[b] Xuedong Lu,[a] Yiping W. Han,[c,d,e]* Yi-Wei Tang[b,h]

Department of Laboratory Medicine, Futian Hospital, Guangdong Medical College, Shenzhen, China[a]; Department of Laboratory Medicine, Memorial Sloan-Kettering Cancer Center, New York, New York, USA[b]; Departments of Periodontics,[c] Pathology,[d] and Reproductive Biology,[e] Case Western Reserve University, Cleveland, Ohio, USA; bioMérieux, Inc., Hazelwood, Missouri, USA[f]; bioMérieux SA, La Balme les Grottes, France[g]; Department of Pathology and Laboratory Medicine, Weill Medical College of Cornell University, New York, New York, USA[h]

We explored the use of matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) for identification of *Fusobacterium nucleatum* subspecies. MALDI-TOF MS spectra of five *F. nucleatum* subspecies (*animalis*, *fusiforme*, *nucleatum*, *polymorphum*, and *vincentii*) were analyzed and divided into four distinct clusters, including subsp. *animalis*, *nucleatum*, *polymorphum*, and *fusiforme*/*vincentii*. MALDI-TOF MS with the modified SARAMIS database further correctly identified 28 of 34 *F. nucleatum* clinical isolates to the subspecies level.

Fusobacterium nucleatum is an opportunistic pathogen, associated with various forms of periodontal diseases and extraoral infections, as well as colorectal cancer (1, 2). *F. nucleatum* is a highly heterogeneous species and was classified into five subspecies: *animalis*, *nucleatum*, *polymorphum*, *vincentii*, and *fusiforme* (3–5). Recent studies based on phylogenetic analysis of the nucleic acid sequences of 16S rRNA, *rpoB*, zinc protease, and 22 other housekeeping genes suggested that *F. nucleatum* subsp. *fusiforme* and *vincentii* be classified into a single subspecies, *F. nucleatum* subsp. *fusiforme*/*vincentii* (4, 5).

Different subspecies may vary in pathogenesis relating to different levels of disease activity (6–8). *Fusobacterium nucleatum* subsp. *nucleatum* is isolated mostly in periodontal diseased sites, whereas *F. nucleatum* subsp. *fusiforme*/*vincentii* is often isolated from healthy sites as normal flora (9). *F. nucleatum* subsp. *animalis* and *polymorphum* are associated with pregnancy complications (7), and *F. nucleatum* subsp. *animalis* is also associated with inflammatory bowel disease (10). Until now, molecular technologies have been the most effective and widely accepted tools for subspecies identification (3, 4, 11). At the subspecies level of *F. nucleatum*, the sequence divergences of 16S rRNA genes were only 0.6% to 1.9%, so full-length sequencing of 16S rRNA was desirable (4).

This study aimed to explore the use of matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) for the identification of *F. nucleatum* at the subspecies level. A commercially available database was amended using 15 *F. nucleatum* isolates comprising the type strains and other well-characterized clinical isolates of the five subspecies and then tested for rapid identification of *F. nucleatum* subspecies.

(This study was presented in part at the 113th General Meeting of the American Society for Microbiology, Denver, CO, 18 to 21 May 2013.)

Bacterial strains and clinical isolates. Five *F. nucleatum* type strains (subsp. *animalis* ATCC 51191^T, subsp. *fusiforme* ATCC 51190^T, subsp. *vincentii* ATCC 49256^T, subsp. *polymorphum* ATCC 10953^T, and subsp. *nucleatum* ATCC 25586^T) and one additional strain of subsp. *nucleatum* ATCC 23726 were obtained from the American Type Culture Collection (ATCC; Manassas, VA). Nine *F. nucleatum* collection strains (two subsp. *animalis*, one subsp. *fusiforme*, one subsp. *vincentii*, three subsp. *polymorphum*, and two subsp. *nucleatum*) were obtained from the Case Western Reserve University collection by phenotypic and genotypic methods as described previously (10). A total of 34 *F. nucleatum* clinical isolates were collected from Case Western Reserve University, Memorial Sloan-Kettering Cancer Center, Harvard University School of Public Health, and bioMérieux, Inc. The isolates were stored in Trypticase soy broth with glycerol at −80°C, cultivated on Columbia agar (Becton, Dickinson and Company, Sparks, MD) supplemented with 5% sheep blood, and incubated at 37°C for 48 h in an anaerobic chamber, as previously described (10).

Full-length 16S rRNA gene sequencing. Bacterial DNA was extracted from a pure colony with a Pure Link genomic DNA minikit (Invitrogen, Carlsbad, CA). The 16S rRNA gene of *Fusobacterium* strains was amplified by PCR using universal primers 17F (5′-GTT TGA TCC TGG CTC AG-3′) and 1512R (5′-TAC CTT GTT ACG ACT T-3′), platinum high-fidelity *Taq* DNA polymerase (Invitrogen), and an Applied Biosystems 2720 thermal cycler (Applied Biosystems, Foster City, CA), as described previously (10). The PCR products were purified and used for DNA sequencing at the Genomics Core Facility (Case Western Reserve

Received 28 January 2015 **Accepted** 29 January 2015
Accepted manuscript posted online 4 February 2015
Citation Nie S, Tian B, Wang X, Pincus DH, Welker M, Gilhuley K, Lu X, Han YW, Tang Y-W. 2015. *Fusobacterium nucleatum* subspecies identification by matrix-assisted laser desorption ionization–time of flight mass spectrometry. J Clin Microbiol 53:1399–1402. doi:10.1128/JCM.00239-15.
Editor: B. A. Forbes
Address correspondence to Yiping W. Han, ywh2102@columbia.edu, or Yi-Wei Tang, tangy@mskcc.org.
* Present address: Yiping W. Han, Division of Periodontics, College of Dental Medicine, and Department of Microbiology & Immunology, College of Physicians and Surgeons, Columbia University Medical Center, New York, New York, USA.
Copyright © 2015, American Society for Microbiology. All Rights Reserved.
doi:10.1128/JCM.00239-15

University, Cleveland, OH) using 17F and 1512R and additional internal primers as described previously (10). The sequences were assembled and aligned using VectorNTI (Invitrogen), and BLAST analysis was performed using NCBI GenBank, Human Oral Microbiome Database, and The Ohio State University core database for subspecies identification.

MALDI-TOF MS. A small portion of one or more colonies of a pure culture was directly smeared onto a FlexiMass disposable target plate (bioMérieux, Marcy l'Etoile, France), immediately followed by addition of 1 μl alpha-cyano-4-hydroxycinnamic acid (CHCA; bioMérieux) matrix solution and allowed to air dry at room temperature. Three independent isolated colonies were deposited for each strain. This procedure, including isolate cultivation, was repeated twice to minimize the effect of variations arising from the process of sample preparation. Spectrum acquisition was performed on a Vitek MS mass spectrometer (bioMérieux) as described previously (12). All spectra were processed by the LaunchPad v2.8 software with baseline correction, peak filtering, and smoothing yielding peak lists with peak apex *m/z* values and corresponding intensities. The peak lists were processed by the Spectral Archiving and Microbial Identification System (SARAMIS) software package (version 4.0.0.14; bioMérieux) for identification.

SARAMIS database amendment. The original SARAMIS database (version 4.10) contained spectra of only one subspecies (subsp. *nucleatum*) and further spectra assigned only to the species level. The six ATCC strains (representing all five subspecies) and nine well-characterized clinical isolates were used to amend the database with *F. nucleatum* subspecies. For each strain, nine spots were analyzed (three different spots of three cultivations), and nine independent spectra were recorded and used to create SuperSpectra for each subspecies using the SARAMIS software (13). As a first step, spectra of individual subspecies were imported and compared by cluster analysis (applying a single linkage agglomerative clustering procedure) and evaluated for consistency. Eventual outliers were reevaluated by comparing individual *m/z* values and noise distribution in raw spectra and by checking for possible sample confusion. When concluding that deviation of a spectrum from the typical pattern was due to a technical error, the spectrum was removed from the database.

In a next step, spectra of all subspecies were compared by cluster analysis, and taxonomic units were defined for which SuperSpectra for automated identification could be computed. At this stage, the low discrimination between the subspecies *vincentii* and *fusiforme* became evident; hence, spectra of both subspecies were used to compute a combined SuperSpectrum. SuperSpectrum computing was performed according to the manufacturer's recommendations in the SARAMIS manual. As *F. nucleatum* subsp. *fusiforme* and *vincentii* are classified into one single subspecies (4, 5), the amended database identified isolates of four subspecies: *animalis*, *fusiforme/vincentii*, *nucleatum*, and *polymorphum*.

MALDI-TOF MS identification. The 34 *F. nucleatum* clinical isolates were analyzed by the Vitek MS system, and subspecies were determined by the above-mentioned *F. nucleatum* subspecies amendment of the database. For each isolate, nine replicate spectra (three different spots of three cultivations) were acquired and matched against the amended SuperSpectra database. SARAMIS generated a confidence-based ranked list of significant matches, with the degree of spectral concordance expressed as confidence levels. By default, a confidence value of ≥80% is considered reliable identification (13). The match with the highest taxonomic precision is ranked first (top hit); for example, 85% confidence for species A subspecies B overrides 95% confidence for species A. Furthermore, any conflicting result with an inconsistent entry on any taxonomic rank (genus, species, subspecies) is highlighted in red to warn the user of a possible mixed-sample result.

Among the five ATCC type strains representing the five *F. nucleatum* subspecies, distinct peak patterns were recorded. In a typical spectrum, 80 to 250 peaks were recorded in the mass range between 3,000 and 20,000 Da, with the highest-intensity peaks consistently positioned in the range of 4,000 to 13,000 Da (Fig. 1A). While the overall similarity of the spectra is evident, several mass peaks allowed the differentiation of subspecies. For example, in the 6,100 to 6,500-Da range, *m/z* 6,176, 6,294, and 6,405 Da were typical for subsp. *polymorphum*; *m/z* 6,148, 6,309, and 6,347 Da were typical for subsp. *nucleatum*; *m/z* 6,150, 6,260, and 6,348 Da for subsp. *vincentii/fusiforme*; and *m/z* 6,177, 6,268, and 6,349 Da were typical for subsp. *animalis* (Fig. 1A). By browsing the genomic sequences of strains ATCC 51191, ATCC 51190, ATCC 23726, ATCC 25586, ATCC 10953, and ATCC 49526 (EMBL-EBI databases as of July 2014; http://www.dkfz.de/srs/), the mass differences can be correlated to variations in protein sequences for some peaks: for example, peaks at *m/z* 6,149 ± 1 and 6,176 ± 1 can be tentatively assigned to ribosomal proteins L33. In subspecies *animalis* and *polymorphum*, an arginine is found at position 41, while a lysine is found in the other subspecies, corresponding to the encountered mass difference of 27 ± 1 Da. Similarly, peaks at *m/z* 6,348 ± 1 and 6,405 ± 1 can tentatively be assigned to ribosomal proteins L32, with the N-terminal methionine cleaved. In subspecies *polymorphum*, a glutamine at position 39 replaces an alanine, as found in the other subspecies.

A total of 135 spectra obtained from the 15 reference strains were retained in the SARAMIS database, representing the five subspecies. A cluster analysis revealed four distinctive clusters (Fig. 1B), with subspecies *fusiforme* and *vincentii* forming a single cluster in accordance with taxonomic studies (4, 5). Although a separation of these two subspecies in separate subclusters is suggested, a high share of background similarity suggested the combination of both subspecies to subspecies *fusiforme/vincentii*. For subspecies *animalis*, the type strain was found to be rather dissimilar to the two clinical isolates but nevertheless included in a subspecies-specific cluster. In subsequent analyses, four subspecies were considered.

A total of 34 clinical isolates were examined in parallel by 16S rRNA gene sequencing and MALDI-TOF MS/SARAMIS. SARAMIS identification of all isolates yielded confidence values of >80% for subspecies, and there was no significant difference among the three independent experiments (data not shown). Twenty-eight (82.4%) of 34 isolates were identified to the subspecies level in accordance with sequence-based identification (Table 1). Among these, 10 belonged to *F. nucleatum* subsp. *animalis*, 8 to *F. nucleatum* subsp. *polymorphum*, 5 to *F. nucleatum* subsp. *nucleatum*, and 5 to *F. nucleatum* subsp. *fusiforme/vincentii*.

Until now, MALDI-TOF MS has had difficulty identifying organisms in the genus *Fusobacterium* to the species level (14), albeit their medical relevance is increasingly recognized (15). For 34 clinical isolates of the validation set, subspecies identification concordant with 16S sequencing was achieved for 28 isolates with the

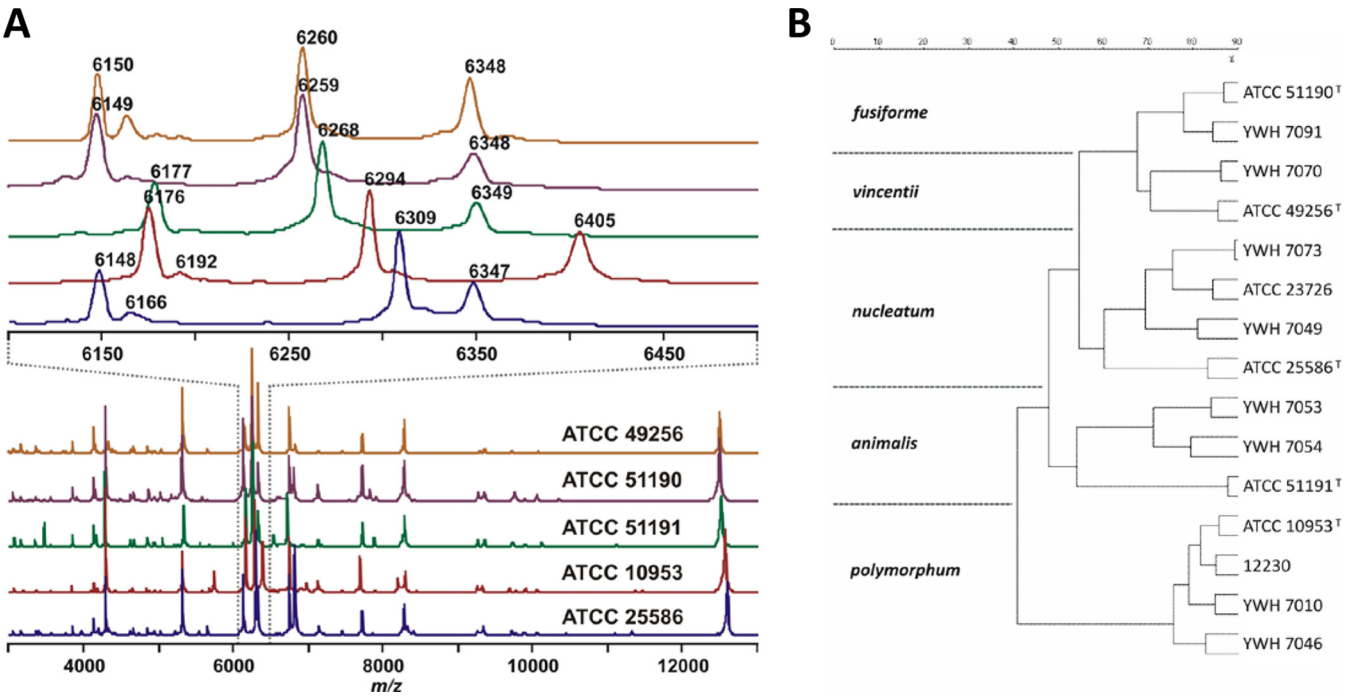

FIG 1 Identification of *F. nucleatum* to subspecies level by MALDI-TOF MS. (A) MALDI-TOF mass spectra of five type strains of the *F. nucleatum* subspecies. The upper plot shows an enlargement (6,100 to 6,500 Da) of the full mass range (3,000 to 13,000 Da), highlighting differences in peak patterns for this range. ATCC 51191, subsp. *animalis*; ATCC 51190, subsp. *fusiforme*; ATCC 49256, subsp. *vincentii*; ATCC 10953, subsp. *polymorphum*; ATCC 25586, subsp. *nucleatum*. (B) Cluster analysis of mass spectra of type and clinical strains of four (five) *Fusobacterium nucleatum* subspecies used for the amendment of the reference database. For each strain, duplicate measurements were considered.

amended SARAMIS database. Discordant subspecies identification was obtained for 4 isolates, while 2 isolates were identified to the species level (Table 1), which may be due to limitations in 16S rRNA sequencing and/or MALDI-TOF technologies. One potential reason for shortcomings in identification performance may be the limited number of isolates available for database expansion on one side and validation on the other side. We are exploring other genomic and proteomic techniques to further study these discordant results.

Further identification will be more accurate after the SARAMIS database is expanded with additional *F. nucleatum* subspecies reference strains. With that, MALDI-TOF MS can rapidly and automatically identify *F. nucleatum* to the subspecies level in a few minutes. As accurate identification of *F. nucleatum* to the subspecies level may improve knowledge about the pathogenicity, epidemiology, and clinical relevance, it is predicted that MALDI-TOF MS can be used in routine practice in clinical microbiology laboratories to identify *F. nucleatum* at the subspecies level.

TABLE 1 *F. nucleatum* subspecies identified by MALDI-TOF MS and 16S rRNA sequencing of 34 clinical isolates

No. of isolates	Subspecies by MALDI-TOF MS	Subspecies by 16S rRNA[a]
10	*animalis*	*animalis*
5	*nucleatum*	*nucleatum*
8	*polymorphum*	*polymorphum*
3	*fusiforme*/*vincentii*	*fusiforme*
2	*fusiforme*/*vincentii*	*vincentii*
1	subsp. not identified	*nucleatum*
1	subsp. not identified	*polymorphum*
1	*polymorphum*	*animalis*
1	*animalis*	*nucleatum*
1	*polymorphum*	*nucleatum*
1	*fusiforme*/*vincentii*	*nucleatum*

[a] 100% similarity was reached for at least one of the sequences in NCBI GenBank, Human Oral Microbiome Database, or The Ohio State University core database.

ACKNOWLEDGMENTS

We thank Wendy Garrett, Erin Thompson, Ying Huang, and Kathy Wilkey for isolate collection and/or technical assistance.

This work was supported in part by a research contract between the Memorial Sloan-Kettering Cancer Center and bioMérieux, Inc. (SK2012-0595) to Y.-W.T. and by National Institutes of Health grants (RO1DE014924, R01CA192111, and RO1DE023332) to Y.W.H.

D.H.P. and M.W. are employees of bioMérieux, Inc., and bioMérieux SA.

REFERENCES

1. **Allen-Vercoe E, Strauss J, Chadee K.** 2011. *Fusobacterium nucleatum*: an emerging gut pathogen? Gut Microbes **2**:294–298. http://dx.doi.org/10.4161/gmic.2.5.18603.
2. **Han YW.** 2015. *Fusobacterium nucleatum*: a commensal-turned pathogen. Curr Opin Microbiol **23C**:141–147. http://dx.doi.org/10.1016/j.mib.2014.11.013.
3. **Strauss J, White A, Ambrose C, McDonald J, Allen-Vercoe E.** 2008. Phenotypic and genotypic analyses of clinical *Fusobacterium nucleatum* and *Fusobacterium periodonticum* isolates from the human gut. Anaerobe **14**:301–309. http://dx.doi.org/10.1016/j.anaerobe.2008.12.003.
4. **Kim HS, Lee DS, Chang YH, Kim MJ, Koh S, Kim J, Seong JH, Song SK, Shin HS, Son JB, Jung MY, Park SN, Yoo SY, Cho KW, Kim DK, Moon S, Kim D, Choi Y, Kim BO, Jang HS, Kim CS, Kim C, Choe SJ, Kook JK.** 2010. Application of *rpoB* and zinc protease gene for use in molecular discrimination of *Fusobacterium nucleatum* subspecies. J Clin Microbiol **48**:545–553. http://dx.doi.org/10.1128/JCM.01631-09.

5. **Kook JK, Park SN, Lim YK, Choi MH, Cho E, Kong SW, Shin Y, Paek J, Chang YH.** 2013. *Fusobacterium nucleatum* subsp. *fusiforme* Gharbia and Shah 1992 is a later synonym of *Fusobacterium nucleatum* subsp. *vincentii* Dzink et al. 1990. Curr Microbiol **66:**414–417. http://dx.doi.org/10.1007/s00284-012-0289-y.
6. **Bolstad AI, Jensen HB, Bakken V.** 1996. Taxonomy, biology, and periodontal aspects of *Fusobacterium nucleatum*. Clin Microbiol Rev **9:**55–71.
7. **Han YW, Wang X.** 2013. Mobile microbiome: oral bacteria in extra-oral infections and inflammation. J Dent Res **92:**485–491. http://dx.doi.org/10.1177/0022034513487559.
8. **Dahya V, Patel J, Wheeler M, Ketsela G.** 13 January 2015. *Fusobacterium nucleatum* endocarditis presenting as liver and brain abscesses in an immunocompetent patient. Am J Med Sci. http://dx.doi.org/10.1097/MAJ.0000000000000388
9. **Gharbia SE, Shah HN, Lawson PA, Haapasalo M.** 1990. Distribution and frequency of *Fusobacterium nucleatum* subspecies in the human oral cavity. Oral Microbiol Immunol **5:**324–327. http://dx.doi.org/10.1111/j.1399-302X.1990.tb00434.x.
10. **Han YW, Shen T, Chung P, Buhimschi IA, Buhimschi CS.** 2009. Uncultivated bacteria as etiologic agents of intra-amniotic inflammation leading to preterm birth. J Clin Microbiol **47:**38–47. http://dx.doi.org/10.1128/JCM.01206-08.
11. **Gonzales-Marin C, Spratt DA, Allaker RP.** 2013. Maternal oral origin of *Fusobacterium nucleatum* in adverse pregnancy outcomes as determined using the 16S-23S rRNA gene intergenic transcribed spacer region. J Med Microbiol **62:**133–144. http://dx.doi.org/10.1099/jmm.0.049452-0.
12. **Hinse D, Vollmer T, Erhard M, Welker M, Moore ER, Kleesiek K, Dreier J.** 2011. Differentiation of species of the *Streptococcus bovis/equinus*-complex by MALDI-TOF mass spectrometry in comparison to *sodA* sequence analyses. Syst Appl Microbiol **34:**52–57. http://dx.doi.org/10.1016/j.syapm.2010.11.010.
13. **Kallow W, Erhard M, Shah HN, Raptakis E, Welker M.** 2010. MALDI-TOF MS and microbial identification: years of experimental development to an established protocol, p 255–276. *In* Shah HN, Gharbia SE(ed), Mass spectrometry for microbial proteomics. Wiley, West Sussex, England.
14. **Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, Raoult D.** 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. http://dx.doi.org/10.1086/600885.
15. **Fedorko DP, Drake SK, Stock F, Murray PR.** 2012. Identification of clinical isolates of anaerobic bacteria using matrix-assisted laser desorption ionization-time of flight mass spectrometry. Eur J Clin Microbiol Infect Dis **31:**2257–2262. http://dx.doi.org/10.1007/s10096-012-1563-4.