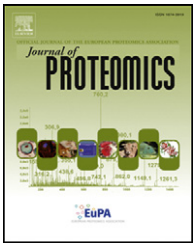
JOURNAL OF PROTEOMICS 98 (2014) 112–122

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot



High natural variability bacteria identification and typing: *Helicobacter pylori* analysis based on peptide mass fingerprinting☆

*Di Xiao, Huifang Zhang, Lihua He, Xianhui Peng, Yandong Wang, Guohui Xue, Peng Su, Jianzhong Zhang**

Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

ARTICLE INFO

Article history:
Received 21 June 2013
Accepted 22 November 2013
Available online 29 December 2013

Keywords:
MALDI-TOF MS
Reference database
Helicobacter pylori
Identification
Typing

ABSTRACT

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) coupled to the original Biotyper database has a poor ability to identify *Helicobacter pylori*. Furthermore, the existing typing methods for *H. pylori* have no obvious correlation with the virulence and pathogenicity of *H. pylori* in East Asia. In this study, MALDI-TOF MS Biotyper system (MBS) was used to identify and type *H. pylori*. In addition, label-free and bioinformatics techniques were used to reveal the protein components of different types of *H. pylori*. A total of 56 *H. pylori* isolates were added to the original reference database. For the 92 *H. pylori* strains validated, the identification efficiency at the species level was improved from 3 (3.2%) to 82 (89.1%) strains. A new ribotyping method for *H. pylori* based on peptide mass fingerprinting was developed. For P1 and P2 type *H. pylori*, respectively, 7 specific peaks at *m/z* 4320, 5202, 5246, 5268, 6066, 6941, and 7128 and 2 specific peaks at *m/z* 5246 and 6941 were found. Between P1 and P2 type strains, 62 proteins were significantly different. A total of 206 proteins for *H. pylori* identification and typing were identified, of which 110 were located on the inner cell membrane and 103 were located in the cytoplasm. The major classifications of these proteins were ribosomal proteins (15.5%) and enzymes (29.6%). MBS is suitable for the identification and typing of variable bacteria such as *H. pylori*, particularly if further super reference spectra are constructed.

Biological significance

Helicobacter pylori (*H. pylori*) possesses very high genetic variability. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) coupled to the original Biotyper system (MBS) has a poor ability to identify *H. pylori* isolated from China. The identification capabilities of MBS for highly variable bacterial species remain to be established. On the other side, Scholars of East Asia and Western dispute the theory that there are obvious correlations between cagA and gastric cancer. The existing typing methods for *H. pylori* based on cagA gene have no obvious correlation with the virulence and pathogenicity of *H. pylori* in East Asia. In light of this phenomenon of Asian enigma, we suppose that there are other genes

☆ Potential competing interests: None.

* *Corresponding author at:* State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, 155 Changbai Road Changping District, Beijing 102206, China. Tel.: +86 1058900754; fax: +86 1058900707.
E-mail address: zhangjianzhong@icdc.cn (J. Zhang).

http://dx.doi.org/10.1016/j.jprot.2013.11.021

beyond cagA that are correlated with the virulence of *H. pylori*. Here, we improved the original database using numerous *H. pylori* isolated from different countries and raised the identification capability of MBS from 3.2% to 89.1%. A new ribotyping method for *H. pylori* based on peptide mass fingerprinting was developed. Furthermore, the protein components of *H. pylori* identification and typing were revealed. These findings thus provide a new way for *H. pylori* identification, typing and the research of pathogenic mechanism.

1. Introduction

As an alternative to biochemical and genome-based identification schemes, direct bacterial profiling using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) is a novel, rapid, and accurate approach for the identification and subtyping of bacterial species. In this technique, specific mass spectra of the mixture of cellular components, primarily proteins and peptides, are obtained directly from whole cells without preliminary separation of cellular components. The identification capabilities of this approach for many clinical and laboratory microorganisms, including Gram-positive and Gram-negative bacteria [1–5], yeast and fungi [6–8], and aerobic, microaerophilic, and anaerobic bacteria [9–11], have been evaluated. *H. pylori* is a Gram-negative, microaerophilic bacterium adapted for survival in the human stomach, where it can cause chronic gastritis and peptic ulcer disease and is an important risk factor for gastric cancer. *H. pylori* has been officially confirmed as a class I carcinogen and remains the only bacterial pathogen that has been established as capable of causing human cancer [12]. *H. pylori* possesses very high genetic variability [13,14]. The identification capabilities of MALDI-TOF MS technology with regard to highly variable bacterial species remain to be established. Ilina et al. [15] recently published an evaluation of MALDI-TOF MS using the Bruker Biotyper system for the identification of *H. pylori*. Their study demonstrated that *H. pylori* can be perfectly identified by the MALDI-TOF MS Biotyper system (MBS, containing 3287 reference spectra, including *H. pylori* 26695 and *H. pylori* J99), and they identified 6 *H. pylori*-specific peaks (*m/z* 4320, 5246, 5529, 5541, and 6066 with frequency greater than 70%, 6946 with a frequency of 59%). However, when we tested the identification capability of this system for *H. pylori* strains isolated from China, none of the *H. pylori* strains could be identified at the species level using MBS (containing 3995 reference spectra, including 7 *H. pylori* strains), and 5 of the *H. pylori*-specific peaks were not specific in our study.

H. pylori is characterized by a large number of gene polymorphisms, and its genome contains obvious regional aggregation. Classification research on *H. pylori* has vital significance for fully understanding the correlation between polymorphisms and the diseases caused by *H. pylori* infection. Currently, the 3 terminal variable regions of cytotoxin-associated gene A (cagA) are regarded as the basis for typing and are used to classify *H. pylori* into East Asia and Western strains [16,17]. Based on the presence or absence of the cagA and vacA genes, *H. pylori* also can be classified as type I and type II strains [18].

The aim of the present study was to add new reference spectra for *H. pylori* isolated from different countries into the Biotyper 2.0 software (Bruker Daltonics, Bremen, Germany), followed by a comprehensive evaluation to validate the identification capability of the Microflex LT-Biotyper system (Bruker Daltonics, Bremen, Germany) for *H. pylori* and to classify the strains of *H. pylori* isolated from different countries according to their peptide mass fingerprints (PMF). Furthermore, the specific peaks and the proteins used to identify and type *H. pylori* were analyzed.

2. Methods

2.1. Bacterial strains and phenotypic identification

A total of 92 *H. pylori* strains from the *H. pylori* strains library of the China CDC were used in this study. The *H. pylori* strains were identified by microscopic examination and urease, catalase, and oxidase activity tests. All strains were grown on Columbia blood agar (CM0331, OXOID) plates and inoculated at 37 °C for 24 to 48 h under microaerophilic conditions.

2.2. Sample preparation and data acquisition

Samples of the examined strains were pre-extracted using previously described procedures [4]. Two spots were prepared for each sample. The instrumentation (Microflex LT, Bruker Daltonics, Bremen, Germany) and data acquisition were operated in accordance with our previous methods (each spectrum contained 100 shots, and 400 shots were superimposed to generate the total spectrum) [4].

2.3. Reference spectra construction and evaluation

A reference database was constructed in accordance with the previous methods [5]. Based on the results of the clustering analysis, these strains, which have common features, were built into the super-reference spectra by the major spectra projection (MSP) function of Biotyper. To evaluate the efficiency of the reference spectra, 56 *H. pylori* strains were used to construct the new reference spectra and were used for cross-validation, and 36 isolates of *H. pylori* isolated from China were used for external validation with the automated option in the Biotyper software.

2.4. Clustering analysis

MSP dendrogram analysis of the *H. pylori* strains was performed as described previously [5,19].

Clustering by the cagA C-terminus: The same primers were used to amplify and sequence the cagA gene of all 56 *H. pylori* strains (forward 5′- ATGACTAACGAAACTATTGA-3′, reverse 5′-TTAAGATTTTTGGAAACCAC-3′). The PCR products of cagA-positive strains were submitted to BGI sequencing analysis.