Analytica Chimica Acta 684 (2011) 8–16

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Typing of unknown microorganisms based on quantitative analysis of fatty acids by mass spectrometry and hierarchical clustering

Tingting Li, Ling Dai, Lun Li, Xuejiao Hu, Linjie Dong, Jianjian Li, Sule Khalfan Salim, Jieying Fu, Hongying Zhong*

Key Laboratory of Pesticides and Chemical Biology, Ministry of Education, College of Chemistry, Central China Normal University, Wuhan, Hubei 430079, PR China

ARTICLE INFO

Article history:
Received 23 August 2010
Received in revised form 27 October 2010
Accepted 28 October 2010
Available online 5 November 2010

Keywords:
Microorganism identification
Fatty acids
Hierarchical clustering
Mass spectrometry
Stable isotope labeling

ABSTRACT

Rapid identification of unknown microorganisms of clinical and agricultural importance is not only critical for accurate diagnosis of infections but also essential for appropriate and prompt treatment. We describe here a rapid method for microorganisms typing based on quantitative analysis of fatty acids by iFAT approach (Isotope-coded Fatty Acid Transmethylation). In this work, lyophilized cell lysates were directly mixed with 0.5 M NaOH solution in d3-methanol and n-hexane. After 1 min of ultrasonication, the top n-hexane layer was combined with a mixture of standard d0-methanol derived fatty acid methylesters with known concentration. Measurement of intensity ratios of d3/d0 labeled fragment ion and molecular ion pairs at the corresponding target fatty acids provides a quantitative basis for hierarchical clustering. In the resultant dendrogram, the Euclidean distance between unknown species and known species quantitatively reveals their differences or shared similarities in fatty acid related pathways. It is of particular interest to apply this method for typing fungal species because fungi has distinguished lipid biosynthetic pathways that have been targeted for lots of drugs or fungicides compared with bacteria and animals. The proposed method has no dependence on the availability of genome or proteome databases. Therefore, it is can be applicable for a broad range of unknown microorganisms or mutant species.

1. Introduction

Rapid identification of microorganisms plays an important role in clinical diagnosis, homeland security, public health, environmental and food safety as well as suitable treatment for plant diseases [1–5]. While Gram stain is still a popular technique to distinguish differences in cellular morphology [6], alternatively phenotypic methods and genotypic methods have been developed for more accurate microbial identification [7,8]. PCR based methods or DNA sequencing [9,10] usually are time-consuming and involve extensive sample preparation. Therefore, considerable efforts have been focused on the invention of new techniques that can provide rapid, sensitive and informative characterization of microorganisms.

Mass spectrometry-based approaches are continuously attracting more and more attention for microorganism identification because of its remarkable speed of data acquisition, high sensitivity of detection and the ability to identify diverse structures of biomarkers provided by tandem MS/MS experiments [11–14]. Typically, mass spectrometry-based approaches are integrated with proteomic strategies. Either top-down protein mass mapping and sequencing [15–18] or bottom-up peptide fragmentation [19,20] in combination with genomic and proteomic database searching has been demonstrated to be useful for reliable microbial identification with sequenced genomes. The major challenge is the identification of unknown microorganisms that do not have complete genome databases, or mutant species with altered genome and proteome, as well as known microorganisms with posttranslationally modified proteins [21]. All of these situations result in un-interpretative mass spectrum. Additionally, the goal of microbial identification is not only the correct grouping of unknown microorganisms with their closest relatives, but also eventually also the exploration of prompt treatment for microbial infections and disease controls. Profiling approaches of both proteomics and genomics cannot provide the activities of enzymes and endpoint output of new or altered genome for unknown or mutant species [22]. Therefore, it is difficult to make judgement whether these new genome or alterations in genome are linked to different regulation of major signal transduction pathways critical for microbial replication and disease infection, consequently leading

* Corresponding author. Tel.: +86 27 67862616.
E-mail address: hyzhong@mail.ccnu.edu.cn (H. Zhong).

0003-2670/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2010.10.047

to the discovery of potentially new targets for therapeutic intervention.

Untargeted metabolite screening offers the possibility to complement genomic and proteomic approaches [23,24]. Although the presence and abundance of specific small metabolites are not directly encoded in the genome, they are indicators of cellular and physiological metabolism and offer a read-out of differences in the underlying signal transduction pathways [25]. Profiling of small metabolites including every kind of lipids, carbohydrates and amino acids has been studied as a molecular signature for the identification of microorganisms [26,27]. But so far, relatively limited successes have been achieved due to the analytical challenges for identification and quantitation of small metabolites with diverse structures.

Microbial identification based on the presence or absence of fatty acids has been used for more than 30 years [28–30]. It has been demonstrated that microbial fatty acid profiles are unique from one species to another [31]. In the present study, quantitative analysis of targeted long chain fatty acids has been investigated for more comprehensive typing of unknown microorganisms. The reasons are summarized in the following: (1) unlike other metabolites, unambiguous identification of targeted long chain fatty acids can be obtained straightforwardly by the combination of both database searching of fragment ions generated by mass spectrometers and comparison of chromatographic retention time with that of standard counterparts that can be commercially available. (2) Because fatty acid identification is usually performed by gas chromatography–mass spectrometric (GC–MS) analysis of methyl esters, there are less interferences from the huge amount of hydrophilic components present in cells that are confronted by LC–MS or MALDI-MS approaches. This GC–MS based detection volatilizes the targeted molecules and concentrates them into an organic phase. (3) This approach also avoids ion suppression resulting from co-existed high abundance components or difficulties in the control of the matrix-to-sample ratio that are usually observed in Matrix Assisted Laser Desorption and Ionization (MALDI) approach. (4) Based on our recently developed iFAT (Isotope-coded Fatty Acid Transmethylation) approach [32,33], a standard mixture of 37 unlabeled fatty acid methyl esters with known concentration can be used as a control. It is then possible to quantitatively determine the content of different fatty acid classes in different cells by measuring the relative intensities of labeled and unlabeled molecular ions and fragment ions. Unknown microorganisms are further classified by tools of computational biology such as hierarchical clustering. Compared with other microbial identification that relies on the presence or absence of targeted molecules, the present approach provides quantitative distances among different species and thus offers quantitative classification of unknown or mutant strains. (5) Fatty acids covalently bonded with other molecules have been implicated in a variety of processes including cellular and physiological energetics [34,35], signaling and metabolism [36,37], cellular membrane formation [38,39] and nuclear receptor activation [40,41]. Thus changes in the level of fatty acids of different organisms can to some extent reveal a possible evolutionary relationship among different species and differences in regulation of major signaling pathways.

In this work, yeast species was used as a model for proof-of-principle demonstration. Clinically and agriculturally, infection with fungi is continuously increasing due to debilitating diseases as well as to the extensive use of broad-spectrum antifungal agents or fungicides [42]. Rapid and accurate identification of fungi is important for predicting the virulence and susceptibilities to antifungal agents or fungicides as well as discovering new therapeutic interventions. In combination with hierarchical clustering [43,44] analysis, quantitatively identified fatty acids in unknown species should make it possible to place them in a dendrogram that can quantitatively reflect their phylogenetic relationship and differences or similarities in fatty acid related pathways.

2. Experimental

2.1. Reagents and apparatus

Yeast cells were purchased from China Center for Type Culture Collection (Wuhan, PR China). n-Hexane (HPLC grade) was purchased from Kermel (Tianjin, China), NaOH and Na_2SO_4 are all of analytical reagent quality and were purchased from Guoyao Co. Ltd., China. Water was obtained from a Milli-Q purification system (Millipore, MA, USA). Supelco^{TM} 37 component FAME mix (standard fatty acid methyl esters) was purchased from Supelco (Bellefonte, PA, USA). Anhydrous methanol, d3-methanol and glass beads are from Sigma-Aldrich (St. Louis, MO, USA). Yeast cell lysis buffer was purchased from Shenergy Biocolor BioScience & Technology, Shanghai, China. Tryptone and yeast extract are from Oxoid (Hampshire, UK).

2.2. Preparation of yeast cell extracts

Yeast cells were grown in YPD media and cells were harvested at different time. Harvested cells were washed three times with PBS buffer before perform lyophilization. Each of 1 mg dried cells were mixed with glass beads (G8772, Sigma) and cell lyses buffer (Shenergy Biocolor BioScience & Technology, Shanghai, China) and then subjected to vortexing on ice for 15 min. Cell lysates of each 1 mg cells were lyophilized and stored in clean glass vials at −20 °C for transmethylation and GC–MS analysis.

2.3. Isotope-coded fatty acid transmethylation (iFAT) of lipids in different yeast cells

In this work, 8 μg of a mixture of 37 standard d0-fatty acid methyl esters with known concentration (10 mg mL^{-1}) were prepared in 800 μL of n-hexane solution to reach a concentration of 10 ng μL^{-1}. Each of 1 mg dry cell lysates stored in glass vials was mixed with 100 μL anhydrous d3-methanol solution of 0.5 M NaOH and 200 μL n-hexane solution. Each vial was treated by 1 min of ultrasonication (100 W, 40 kHz). Then let the vials stand on ice for 5 min before the supernatant hexane parts were transferred. Combine equal volumes of n-hexane solution of standard d0-fatty acid methyl esters with those derived by d3-methanol. The combined n-hexane layers were then washed with pure water three times in order to remove trace methanol and NaOH. Anhydrous Na_2SO_4 was used to absorb trace water remaining in the hexane solution.

2.4. GC–MS analysis and identification of fatty acid methyl esters

1 μL n-hexane solution of d3- and d0-differentially labeled fatty acid methyl esters that were prepared by the previously described procedure was separated by capillary gas chromatography (Rtx-5MS, 0.25 μm in thickness, 30 m × 0.25 mm i.d., USA) and analyzed by a quadruple mass spectrometer (GC–MS QP2010PLUS, Shimadzu, Japan). The column oven temperature was programmed to reach 250 °C during a program of 45 min. Within the whole process of analysis, the injector temperature was kept at 220 °C. Splitless injection was performed with 1 min sampling time. The interface temperature was set at 270 °C and the ion source was maintained at 200 °C. 70 eV ionization energy of impact ionization (EI) was used to fragment the eluent from capillary GC. Mass spectra were recorded in full scan mode with mass-to-charge ratio (m/z) ranging from 50 to 600 unit.

Unambiguous identification of fatty acid methyl esters was achieved by both database searching and comparison of chro-