Journal of Microbiological Methods 86 (2011) 327–336

Contents lists available at ScienceDirect

Journal of Microbiological Methods

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

Evaluation of MALDI-TOF MS as a tool for high-throughput dereplication

Jonas Ghyselinck [a,*], Koenraad Van Hoorde [a,b], Bart Hoste [a], Kim Heylen [a], Paul De Vos [a]

[a] Laboratory of Microbiology, Department of Biochemistry and Microbiology, Ghent University, K.L. Ledeganckstraat 35, B-9000 Gent, Belgium
[b] Faculty of Applied Engineering Sciences, University College Ghent, Schoonmeersstraat 52, B-9000 Gent, Belgium

ARTICLE INFO

Article history:
Received 4 February 2011
Received in revised form 31 May 2011
Accepted 5 June 2011
Available online 13 June 2011

Keywords:
Dereplication
Diversity
MALDI-TOF
rep-PCR

ABSTRACT

The present study examined the suitability of matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) for the rapid grouping of bacterial isolates, i.e. dereplication. Dereplication is important in large-scale isolation campaigns and screening programs since it can significantly reduce labor intensity, time and costs in further downstream analyses. Still, current dereplication techniques are time consuming and costly. MALDI-TOF MS is an attractive tool since it performs fast and cheap analyses with the potential of automation. However, its taxonomic resolution for a broad diversity of bacteria remains largely unknown. To verify the suitability of MALDI-TOF MS for dereplication, a total of 249 unidentified bacterial isolates retrieved from the rhizosphere of potato plants, were analyzed with both MALDI-TOF MS and repetitive element sequence based polymerase chain reaction (rep-PCR). The latter technique was used as a benchmark. Cluster analysis and inspection of the profiles showed that for 204 isolates (82%) the taxonomic resolution of both techniques was comparable, while for 45 isolates (18%) one of both techniques had a higher taxonomic resolution. Additionally, 16S rRNA gene sequence analysis was performed on all members of each delineated cluster to gain insight in the identity and sequence similarity between members in each cluster. MALDI-TOF MS proved to have higher reproducibility than rep-PCR and seemed to be more promising with respect to high-throughput analyses, automation, and time and cost efficiency. Its taxonomic resolution was situated at the species to strain level. The present study demonstrated that MALDI-TOF MS is a powerful tool for dereplication.

1. Introduction

The last decade has seen a renewed interest in bacterial cultivation. New approaches for isolating bacteria have been developed, either through adjusting growth and incubation conditions using alternative gelling agents such as gellan gum (Tamaki et al., 2009) and prolonging incubation times (Davis et al., 2005), or through elaborating new technologies such as high-throughput culturing methods (Connon and Giovannoni, 2002) and diffusion growth chambers (Kaeberlein et al., 2002). To significantly reduce the work load once bacteria are obtained in culture, most studies perform a dereplication step. Conventionally, dereplication refers to the process of eliminating knowns from unknowns. However, the term is often used in an unconventional way, referring to the process of recognizing identical isolates at a specific taxonomic level and grouping them accordingly. Subsequent selection of representatives of each group reduces the number of isolates to be analyzed in further downstream analyses, and thus prevents unnecessary screening efforts. Dereplication in this meaning originally referred to the grouping of bacterial isolates at the lowest taxonomic level, the strain level (Bull et al., 1992). However, nowadays the term is somewhat ambiguous and often used in a broader sense, also indicating grouping at subspecies (De Clerck and De Vos, 2002), species or any higher taxonomic level (Coorevits et al., 2008; Heyrman and Swings, 2001). A broad range of techniques has been used in the context of dereplication such as repetitive element sequence based Polymerase Chain Reaction (rep-PCR) (De Clerck and De Vos, 2002), randomly amplified polymorphic DNA (RAPD) (Martin-Platero et al., 2009) and fatty acid methyl ester (FAME) analyses (Coorevits et al., 2008). In some studies, techniques have been used of which the taxonomic resolution was not validated, e.g. Boroczky et al. (2006) used a specific type of GC-analysis for dereplication of a set of bacteria isolated from marine environments. A suitable dereplication technique should comply with the following criteria: (i) hold a universal character, i.e. applicability to all bacterial strains; (ii) robustness; (iii) produce easy to interpret data; (iv) have a high taxonomic resolution and (v) provide the possibility of high-throughput application/automation with low operational costs and labor intensity.

Rep-PCR has proven to be a powerful tool in microbial ecology and environmental microbiology (Ishii and Sadowsky, 2009). It is a widely applied DNA fingerprinting technique targeting repetitive sequences interspersed throughout the bacterial genome (Kang and Dunne, 2003), and largely fulfills above-mentioned criteria that make a

* Corresponding author at: Laboratory of Microbiology, Department of Biochemistry and Microbiology, Ghent University, K.L. Ledeganckstraat 35, B-9000 Gent, Belgium. Tel.: +32 92645101; Fax: +32 92645092.
E-mail address: jonas.ghyselinck@ugent.be (J. Ghyselinck).

0167-7012/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.mimet.2011.06.004

technique suitable for dereplication. PCR amplification of the DNA between these repetitive elements and subsequent electrophoresis results in easy to interpret bacterial fingerprints that allow differentiation at the subspecies to strain level for a wide range of bacterial species (Dombek et al., 2000; Olive and Bean, 1999; Rademaker et al., 2008; Souza et al., 2010). Furthermore, rep-PCR is a robust technique since factors like e.g. culture age and the number of subcultures prior to DNA extraction have shown not to influence the genomic fingerprint (Kang and Dunne, 2003). Nevertheless, this technique also has drawbacks. Trials with different primer sets may be required to produce good quality fingerprints (Masco et al., 2003) and both intra- (Rasschaert et al., 2005) and interlaboratory reproducibility (Deplano et al., 2000) can sometimes be lacking. These shortcomings hamper (semi-) automation of the technique, making rep-PCR quite laborious to be used as dereplication tool in extensive isolation campaigns.

Matrix assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) is a more recent technique in microbiology and has become an important tool with promising applications, especially in diagnostics (Bizzini and Greub, 2010; Eigner et al., 2009; Murray, 2010). MALDI-TOF MS generates protein mass spectra which can be used to group and identify bacteria. These mass spectra contain mainly peaks corresponding to ribosomal proteins because of their very high abundance in the bacterial cell (Ryzhov and Fenselau, 2001). It could be an interesting tool for dereplication since it has the same advantages as rep-PCR (i.e. applicability for a wide range of bacterial species, generation of easy interpretable data and robustness) with the additional plus-point that it could be automated, resulting in time and cost reduction. Furthermore, the reagents required to prepare bacterial cell extracts and to do the analysis are cost effective. Numerous studies have already explored whether MALDI-TOF MS has the ability to discriminate at the strain level (Arnold and Reilly, 1998; Dickinson et al., 2004; Siegrist et al., 2007; Vargha et al., 2006). However, these studies were systematically limited to a specific taxon (particular genus or species), and therefore insufficient to evaluate the taxonomic resolution of MALDI-TOF MS as a broad range dereplication tool.

In this study, the applicability of MALDI-TOF MS for high-throughput dereplication of a large and unidentified variety of bacterial isolates that were isolated from the potato rhizosphere in Peru and Bolivia was evaluated. Rep-PCR was performed in parallel as a benchmark and both techniques were compared based on grouping of isolates, taxonomic resolution, reproducibility, suitability for high-throughput automation and time and cost effectiveness.

2. Materials and methods

2.1. Bacterial isolates

Bacteria used in this study were isolated from the rhizosphere of potato plants from the Central Andean Highlands. In short, 5 ml phosphate-buffered saline and 10 sterile glass beads (6 mm) were added to 1 g rhizosphere soil, and vortexed for 2 min. Serial dilutions (10^0–10^{-6}) were plated (100 µl) on ten-fold diluted Trypticase Soy Agar (TSA), supplemented with 0.005% (w/v) cycloheximide to inhibit fungal growth. After incubation for 48 h at 28 °C, isolates were picked and subcultured to purity.

2.2. Rep-fingerprinting

Genomic DNA was released from the bacterial cells through alkaline lysis. Therefore, a small amount of cells was lysed in 20 µl alkaline lysis buffer (0.25% (w/v) SDS and 0.05 M NaOH) for 15 min at 95 °C. Subsequently, 180 µl sterile milliQ-water was added and lysates were immediately used for PCR. Rep-PCR was performed with the $(GTG)_5$-primer because in-house experience showed this primer targeted the largest bacterial diversity (unpublished data). The PCR-mixture was prepared as described previously (Rademaker and de Bruijn, 1997). Amplification was performed in a GeneAmp PCR System 9600 (Applied Biosystems) with the following temperature-time profile: 7 min 95 °C, 30 cycles of 94 °C for 1 min, 40 °C for 1 min and 65 °C for 8 min, and a final step of 16 min at 65 °C. Electrophoresis was performed in 1×TAE buffer (40 mM Tris, 20 mM acetic acid and 1 mM EDTA) on a 1.5% agarose gel (w/v) under highly standardized conditions (55 V, 400 mA, 960 min, 4 °C). Fourteen samples were loaded per gel. Four reference markers, 6 µl each composed of 1.10 µl Molecular Ruler 500 bp (Bio-Rad), 1.40 µl Molecular Ruler 100 bp (Bio-Rad), 2 µl TE buffer (1 mM EDTA, 10 mM Tris–HCl (pH 8.0)) and 1.50 µl loading dye, were included on every gel. Profiles were visualized under ultraviolet light after staining with ethidium bromide. Digitized images of gels were normalized and analyzed with the BioNumerics 5.1 software (Applied Maths, Belgium). Similarity matrices of densitometric curves of the gel tracks were calculated with Pearson's product-moment correlation coefficient. Cluster analyses of similarity matrices were performed by unweighted pair group method with arithmetic averages (UPGMA). Reproducibility was assessed by analyzing a random subset comprising 10% of all isolates (24 out of 249) in triplicate (starting from growth and DNA extraction to analysis of fingerprint).

2.3. MALDI-TOF MS

2.3.1. Preparation of cell extracts

Isolates were grown from stock on tenfold diluted TSA for 48 h at 28 °C and subcultured twice prior to analysis to ensure all isolates were in the same physiological state. For preparation of the extracts, a small amount of bacterial cells was picked up and suspended in 300 µl milliQ water. Next, 900 µl of absolute ethanol was added and the bacterial suspension was centrifuged for 3 min at 18,000×*g*. After removing the supernatant, the bacterial pellet was resuspended in 50 µl formic acid (70%). Finally, 50 µl of acetonitrile was added and mixed until complete suspension. The extract was centrifuged for 3 min at 18,000×*g* and the supernatant was used for MALDI-TOF MS analysis or was preserved at −20 °C for later use.

2.3.2. MALDI-TOF MS analysis

Bacterial cell extracts (1 µl) were spotted on a 384 Opti-TOF 123 mm×81 mm stainless steel MALDI-TOF MS target plate (AB Sciex) and dried at room temperature. Subsequently, the sample spot was overlaid with 1 µl of a 0.5% (w/v) α-cyano-4-hydroxycinnamic acid (α-CHCA) in 50:48:2 acetonitrile:water:trifluoroacetic acid solution. The spotted plate was analyzed with the 4800 *Plus* MALDI TOF/TOF Analyzer (AB Sciex) which was used in linear, positive-ion mode. Ions were generated by a 200 Hz tripled UV Nd:YAG laser, accelerated at 20 kV through a grid at 19.2 kV and separated according to their *m/z* ratio in a 1.5 m long linear, field-free drift region. Each generated spectrum resulted from 40 laser shots at 50 random positions within the measuring spot. MALDI-TOF mass spectra were generated in the mass range 2–20 kDa. Calibration was performed with the Protein Calibration Standard I (Bruker) (composition: insulin ($[M+H]^+$, *m/z* 5734.6), ubiquitin I ($[M+H]^+$, *m/z* 8565.9), cytochrome C ($[M+H]^+$, *m/z* 12361.5), myoglobin ($[M+H]^+$, *m/z* 16952.3)) to which ACTH Fragment 18–39 MALDI-MS Standard ($[M+H]^+$, *m/z* 2465.7) (Sigma-Aldrich) was added. With every set of measurements, the Bruker Bacterial Test Standard (Bruker) was included as a positive control.

2.3.3. Analysis of spectral data

Mass spectra were obtained in t2d format and were converted to txt files using the Data Explorer 4.9 software (AB Sciex). The txt files were imported in BioNumerics 5.1 software (Applied Maths, Belgium) and converted to fingerprints for further analyses. To obtain reliable data analysis, the spectra with extensive noise and/or insufficient