Antonie van Leeuwenhoek (2013) 103:603–615
DOI 10.1007/s10482-012-9844-6

ORIGINAL PAPER

Large scale MALDI-TOF MS based taxa identification to identify novel pigment producers in a marine bacterial culture collection

Marit H. Stafsnes · Marius Dybwad · Anders Brunsvik · Per Bruheim

Received: 14 September 2012 / Accepted: 30 October 2012 / Published online: 7 November 2012

Abstract A challenge in the rational exploitation of microbial culture collections is to avoid superfluous testing of replicas. MALDI-TOF MS has been shown to be an efficient dereplication tool as it can be used to discriminate between bacterial isolates at the species level. A bacterial culture collection of more than 10,000 heterotrophic marine bacterial isolates from sea-water surface layers of the Norwegian Trondheimsfjord and neighbouring coastal areas has been established. A sub-collection of pigmented isolates was earlier screened for novel carotenoids with UVA-Blue light absorbing properties. This was a comprehensive analytical task and it was observed that a significant number of extracts with identical pigment profile were recovered. Hence, this study was undertaken to explore the use of MALDI-TOF MS as a dereplication tool to quickly characterize the bacterial collection. Furthermore, LC-DAD-MS analysis of pigment profiles was performed to check if pigment profile diversity was maintained among isolates kept after the potential MALDI-TOF MS selection step. Four hundred isolates comprising both pigmented and non-pigmented isolates were used for this study. The resulting MALDI-TOF MS dendrogram clearly identified a diversity of different taxa and these were supported by the pigment profile clustering, thus linking the pigment production as species-specific properties. Although one exception was found, it can be concluded that MALDI-TOF MS dereplication is a promising pre-screening tool for more efficient screening of microbial culture collection containing pigments with potential novel properties.

Keywords Marine bioprospecting · bacteria · Maldi-ToF MS dereplication · pigmentation · LC-MS

M. H. Stafsnes · M. Dybwad · P. Bruheim (✉)
Department of Biotechnology, Norwegian University of Science and Technology, Sem Sælands vei 6/8, 7491 Trondheim, Norway
e-mail: Per.Bruheim@ntnu.no

M. Dybwad
Norwegian Defence Research Establishment, P.O. Box 25, 2027 Kjeller, Norway

A. Brunsvik
Department of Biotechnology, SINTEF Materials and Chemistry, Sem Sælands vei 2a, 7465 Trondheim, Norway

Introduction

Microorganisms are an immense resource of diverse natural products that are candidates for drug development, food and feed additives, and other industrial products. One major challenge to natural product drug research is that it is labour intensive to isolate large bacterial collections and fully characterise active compounds from extracts, such that it requires a great deal of effort to produce adequate quantities of active compounds. The ability to quickly and reliably differentiate among related bacteria and thereby

identify identical isolates is essential for efficient screening to prevent unnecessary work (Klinkenberg et al. 2011). For example, earlier it was found that the candicidin biosynthesis gene cluster was widely distributed among *Streptomyces* spp. both from marine sediments and surface layers and was repeatedly identified during a search for novel bioactive secondary metabolites (Jørgensen et al. 2009). The process to avoid such waste of resources is defined as dereplication. The selection of isolates to be further processed can be based on their phenotypic qualities or genetic differences or on the production of secondary metabolites. Matrix-assisted laser desorption/ionization-Time-of-flight mass spectrometry (MALDI-TOF MS) has become an important tool in bacterial identification and species differentiation. Hence, it can also be evaluated for use as a dereplication tool (Ghyselinck et al. 2011). MALDI-TOF MS can be used for fast differentiation of isolated cultivable bacteria based on their mass fingerprints (phyloproteomics) in the 2–20 kDa area. In the last decade MALDI-TOF MS has been increasingly studied for potential application to identification and typing of microorganisms. Very recently, MALDI-TOF MS has been introduced in clinical routine microbiological diagnostics with marked success (Welker and Moore 2011; Elssner et al. 2011). In the literature, there are also several examples of the use of MALDI-TOF MS in taxonomic studies at species and subspecies level (Dieckmann et al. 2008; Dubois et al. 2010; Rezzonico et al. 2010; Verroken et al. 2010) and also a few other applications such as screening alkylphenol polyethoxylate-degrading bacteria (Ichiki et al. 2008) and seafood spoilage and pathogenic Gram-negative bacteria (Bohme et al. 2010b). However, few studies have been reported when characterizing the total bacterial population in larger environmental studies (Dieckmann et al. 2005; Ghyselinck et al. 2011; Koubek et al. 2012; Dybwad et al. 2012). To our knowledge, MALDI-TOF MS has not been used to describe such a large and heterogeneous bacterial collection as the one presented in this report.

Marine ecosystems are presumed to be more heterogeneous than soil ecosystems at the bacterial level (Tringe et al. 2004) and currently The World Register of Marine species (WoRMS) contains 3,684 marine bacterial species (Appeltans et al. 2012). The commercial databases for bacterial identification, SARAMIS and Biotyper, contain MS-spectra from about 2,000 different bacteria species (AgnosTec 2012; Daltonics 2012) with the emphasis on validating clinical isolates. The identification frequency can therefore not be expected to be very high for marine environmental isolates. However, a distance-related dendrogram calculated from the mass fingerprints of the bacterial isolates analysed can be constructed and used to evaluate a bacterial culture collection. Ideally it would be useful if the distance score level could be used to classify isolates at the genus, species and strain levels. However, it has been thoroughly verified that the distance that separates species varies significantly. Thus, it is hard to define a strict cut-off value for isolates grouped together below a node in the dendrogram to the same species (Welker and Moore 2011). For example, Verroken and co-workers reported that the distance level that defined different *Nocardia* species varied from 150 to 650 (Verroken et al. 2010). Irrespectively, for the purpose of using MALDI-TOF MS as a dereplication tool this cannot be seen as a disadvantage as the technology can rapidly reveal cryptic species among large numbers of related isolates in large environmental cultural collections. Therefore, the cut-off value for species discrimination is not so critical (Welker and Moore 2011), i.e. unique isolates will be identified irrespective of whether the node is at distance level of 650 or 200 to the nearest neighbouring cluster (Veloo et al. 2011; Sauer et al. 2008). However, as a rule of thumb it is accepted that isolates with distance levels below 500 belong to the same species and distance levels higher than this do not give much information about the relationship (i.e. at the genus level and higher) between the isolates.

In a previous study to characterize marine pigmented bacteria and to identify novel carotenoids with UVA-Blue light absorbing properties (Stafsnes et al. 2010), we experienced a heavy work-load by using LC-DAD-MS pigmentation analysis as the only selection criteria. Therefore, it is of interest to develop a more rational screening tool for this type of study. Although rapid high-throughput screening protocols for LC–MS analysis have been developed (Stafsnes et al. 2010), the cultivation and extraction steps add to the total work load. It was also observed that many isolates produced the same pigmentation profile and in two cases it was confirmed that isolates producing the same pigmentation profile belong to the same bacterial species (Stafsnes et al. 2010). Algae have been proven to contain division- or class-specific carotenoids and

the production of carotenoids can therefore be used as a taxonomic marker (Takaichi 2011). For marine bacteria the same relationship has not been found for general application; species from different families can produce the same carotenoid (e.g. the ubiquitously produced zeaxanthin). Prodigiosin was first believed to be a genus specific pigment (*Serratia* sp.) (Lewis and Corpe 1964; Rettori and Durán 1998) but to date it has been found in a variety of genera. Flexirubin can be considered a taxonomic marker for *Flavobacteria* isolated from soil or freshwater but the same has not been proven for marine species (Reichenbach et al. 1980). On the other hand, a clear division of pigmented and non-pigmented genera in the same family, e.g. *Flavobacteriaceae* (Bae et al. 2007), or species in the same genus, e.g. *Pseudoalteromonas* (Vynne et al. 2011), has been found, thus indicating that the ability to produce pigmented colonies to some extent can be used in taxonomic studies. Most pigmented bacteria produce between three and five different pigments thus having a pigment profile rather than a single pigment for comparison, thereby having considerable potential value in the taxonomy and ecology of microbial communities in natural environments. More broadly, secondary metabolites as phenotypic markers in fungi has been presented in a review by Larsen and co-workers (Larsen et al. 2005). They conclude that the profile of biosynthetic families is always species specific, while individual metabolite biosynthetic families have been found in both phylogenetically closely related and distantly related species. Also secondary metabolites produced by marine invertebrate-associated bacteria have been demonstrated to be valid as taxonomic markers (Hou et al. 2012). With the use of LC/MS-principal component analysis they showed that this metabolomics tool can efficiently assist with strain dereplication, distinguish taxonomically identical species and used to discover new natural products. Image analysis, based on macromorphological phenotypic characters, has also been tested as a dereplication tool on fungal cultures (Hansen et al. 2003; Hansen and Carstensen 2004; Andersen et al. 2005).

This study was initiated to evaluate if a rapid MALDI-TOF MS dereplication step can provide considerable rationalisation by drastically limiting the number of isolates to be analysed by LC-DAD-MS. Importantly, it must be verified that pigmentation diversity is preserved during the MALDI-TOF MS selection step. Also, due to the high number of both pigmented and non-pigmented isolates available in our microbial culture collection, the study should provide information on pigmentation diversity versus bacterial diversity of general ecological interest.

Materials and methods

Cultivation

Cultivations for production of biomass for characterization of the culture collection were performed on agar plates and in shake flasks using YPS medium as described earlier (Stafsnes et al. 2010). YPS-agar plates were stamped from frozen 96-well plates or streaked from the frozen glycerol stocks from the culture collection and kept at 25 °C for 8 days. When pure culture and pigment production was ascertained, single colonies were picked and used to inoculate the shake flask for biomass production. For MALDI-TOF MS analysis, agar plates were reinoculated and cultivated for 2–3 days ensuring sufficient growth and fresh cultures.

Shake flask cultivations were performed in 250 mL shake flasks filled with 50 mL YPS medium. They were incubated for 3–6 days at 25 °C in a Minitron shaking incubator (orbital 25 mm) operated at 200 rpm. They were then harvested by centrifugation at 7,000 rpm (13,900 g), washed with water and re-centrifuged. The cell pellets were frozen and stored at −20 °C prior to extraction of pigments.

Extraction of pigments

The washed cell pellets were frozen at −20 °C to facilitate extraction with organic solvents. 1 mL of a mixture of methanol and acetone (7:3) was added to the cell pellets from shake flasks and mechanically mixed and vortexed for 30 s to ensure even dispersal of the biomass. The suspensions were heated in a water bath at 55 °C, vortexed for 10 s and centrifuged (3200×g, 10 min, 4 °C). If the supernatant was pale and/or the pellet still coloured the extraction step was repeated after removal of the supernatant. Further treatments such as additional cell disintegration with lysozyme were not performed for low yielding extractions as the focus was high throughput screening and not quantification of the pigment production.

LC-DAD-MS analysis

LC–MS analyses were performed on an Agilent Single Quadrupol SL mass spectrometer equipped with an Agilent 1100 series HPLC system. The LC system was equipped with an auto-sampler injecting 10 µL sample and a diode array detector (DAD) that recorded UV/VIS spectra in the range from 200 to 650 nm. Analytes were ionized using a chemical ionization source with settings of 325 °C dry temperature, 350 °C vaporizer temperature, 50 psi nebulizer pressure and 5.0 L/min dry gas. The MS was operated in single ion monitoring (Simmons et al. 2008.) mode for 90 % of cycle time and in scan mode for 10 % of the cycle time. 21 carotenoid masses ($M + H^+$) were included in the SIM list (Stafsnes et al. 2010). Separations were done on a Zorbax rapid resolution cartridge RP C_{18} column with dimensions 2.1 × 30 mm, with column temperature 40 °C and using the Mobile phases A: methanol–water (80:20) and B: dichloromethane. The pigments were eluted with an increasing percentage of mobile phase B using the following gradient: 10 % B from 0 to 0.5 min, then a linear gradient to 70 % B until 10 min. Column flow was kept at 0.4 mL/min. ß-carotene (Sigma) and astaxanthin (Sigma) were used as external standards. Variation of retention times was accounted for by running reference samples after every 10 injections. Samples and standards were filtered through a syringe polypropylene filter (0.2 µm, Pall Gelman) and stored in amber glass vessels at −80 °C under N_2 atmosphere. All solvents and water were HPLC grade and chemicals were analytical grade.

A full manual pigment matrix was made of the HPLC data containing all recognizable peaks/UV–Vis spectra from all HPLC chromatograms. Drift in the elution time of the pigments was manually accounted for on the basis of the retention times of the standards used. The matrix consisted of isolates versus peak (with unique Rt, UV–Vis and *m/z* values) with binary values and was made manually by scoring each metabolite as present or absent in each individual sample. This full manual metabolite matrix consisted of 82 x-variables (known and unknown pigments) and 97 y-variables (bacterial isolates) for the objects. The matrix was analysed with Hierarchical Average-linkage using Euclidean distance with Unscrambler X software, version 10.0.1 (CAMO ASA, Oslo).

MALDI-TOF MS

Sample preparation

A standard protein extraction protocol was used (Sauer et al. 2008; Marklein et al. 2009). In brief, a small amount of bacterial cells was picked (one 1 µL sterile loop) and inactivated with 1 mL 80 % ethanol in water. After centrifugation and removal of the supernatant, cell pellets were extracted with 60 µL of a acetonitrile/formic acid/water (50:35:15, v/v) mixture by vortexing for 1 min. The extract was centrifuged for 2 min at 18,000×g and the supernatant was used directly for MALDI-TOF MS analysis or stored at −20 °C for later use. The supernatant was deposited onto MSP 96 ground steel targets (Bruker Daltonik, Bremen, Germany) as triplicates at a volume of 1 µL and, after drying at room temperature, overlaid with 1 µL of saturated alpha-cyano-4-hydroxycinnamic acid (CHCA) matrix solution in acetonitrile: water: TFA (50: 47.5: 2.5, v/v).

CHCA was obtained from Bruker Daltonik. Ethanol was from Sigma–Aldrich (Steinheim, Germany). Trifluoroacetic acid and acetonitrile were from Merck (Darmstadt, Germany), formic acid from Riedel de Haën (Seelse, Germany). Water was prepared on a Milli-Q plus 185 apparatus (Millipore, Bilerica, MA, USA). All chemicals were of analytical grade.

MALDI-TOF MS analysis

Mass spectra were acquired using a Microflex MALDI-TOF MS instrument (Bruker Daltonik) in linear positive mode. The instrument was operated using Flex Control 3.0 (Bruker Daltonik) and the MALDI Biotyper 2.0 microbial identification software (Bruker Daltonik) in real-time classification mode with default settings (mass range 2,000–20,000 *m/z*; ion source 1 (IS1) 20 kV; ion source 2 (IS2) 17 kV; lens voltage 7 kV, pulsed ion extraction 150 ns). The laser intensity was automatically adjusted using the fuzzy control option in Auto Execute mode. The whole process from MALDI-TOF MS measurement to identification was performed automatically without any user intervention. The reference database used was the entire Bruker Taxonomy database (v3.1.1.0 containing 3,740 library entries). External calibration of the instrument was performed using the Bacterial Test Standard (BTS, #255343, Bruker Daltonik). A sum mass spectrum was

generated from each target spot by summing up 240 individual laser shots. The software provides a log score and a cut-off log score of 2 was used to validate identification at the species level, as recommended by the manufacturer. The Biotyper 2.0 software was also used to generate a main spectra library (MSP) entry for each measured bacterial isolate (Sauer et al. 2008). To generate a single MSP for each bacterial isolate, the sum mass spectra were baseline subtracted, smoothed and processed with the MSP-creation tool available in the Biotyper 2.0 software using default settings. A MSP-based dendrogram was further constructed with the dendrogram tool embedded in the Biotyper 2.0 software using default clustering settings and the complete linkage algorithm.

A number of marine strains did not give any spectra with the default settings in the Biotyper 2.0 software and therefore manual acquisition of the spectra was performed to be able to see how these isolates clustered together. The open-source software MoverZ by Genomic Solutions® was used for mass fingerprint generation and SPECLUST, a web tool available at http://bioinfo.thep.lu.se/speclust.html (Alm et al. 2006), was used for hierarchical clustering and dendrogram calculation.

Results and discussion

Cultivation and selection of isolates from the culture collection for MALDI-TOF MS analysis

The culture collection of heterotrophic marine bacteria established in 2004/2005 contained originally over 10,000 primary isolates sampled from the Norwegian Trondheimsfjord and neighbouring coastal areas (Stafsnes et al. 2010). The pigmented sub-collection was established by a second cultivation stage by picking pigmented primary isolates and this collection consisted of approximately 2,000 isolates. Both collections were grown using both liquid and solid YPS media in 96-well plate format. However, during establishment of the medium and other cultivation conditions it was frequently, and not unexpectedly, observed that some primary isolates lost viability and that some secondary isolates also lost pigmentation even though cultivated using the same conditions as for the primary stage. Figure 1 shows a typical well plate from primary stage (a) and the secondary stage (b) enriched with pigmented strains, exemplifying and visualizing the general observations.

Eight 96-well plates (approximately 750 isolates) from the pigmented sub-collection were selected for this study. In addition, seven 96-well plates from the original primary collection were included as total population samples. The survival of these isolates was satisfactorily high; 73–90 % with an average of 80 % per plate. 5–25 % of the pigmented isolates per plate had lost their ability to produce pigments after thawing and the first round of cultivation. Two-hundred and fifty of these isolates (115 pigmented and 135 non-pigmented) were randomly chosen for MALDI-TOF MS analysis. Furthermore, 150 pigmented isolates from a previous pigmentation analysis study were included (Stafsnes et al. 2010). The reasons for including these isolates were that they were isolated from other sampling sites and that they had already been thoroughly characterized with regard to pigmentation production. They had also been through several rounds of repeated cultivation on solid media to verify pure culture status. Ten of the isolates had been identified with 16S rRNA analysis.

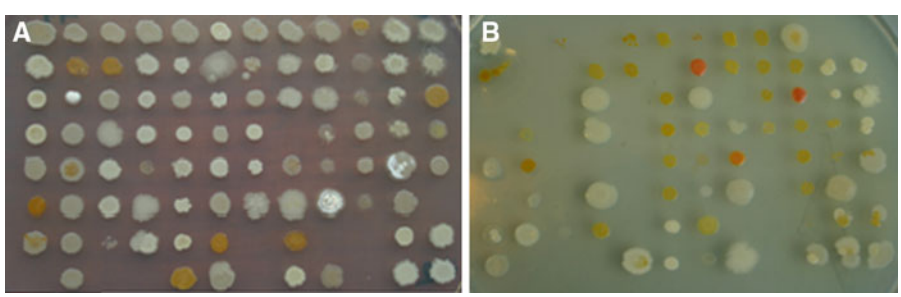


Fig. 1 Marine isolates grown on YPS medium for 8 days. **a** total population **b** pigmented colonies

MALDI-TOF MS fingerprinting and classification of bacterial isolates

MALDI-TOF MS spectra were generated from protein extracts from the 400 selected isolates with three technical replicates, and 373 of these spectra passed the default quality criteria of the Biotyper 2.0 software. Several studies have concluded that the direct smear method is to be preferred due to the much shorter preparation time than with chemical extraction. The marginally higher identification scores obtained with the extraction protocol does not justify the extra preparation time (van Veen et al. 2010; Fournier et al. 2012; Bizzini et al. 2011; Haigh et al. 2011). This is of particular interest within clinical diagnostics. However, for studies on environmental isolates where growth rates, and thereby sampling time points, can vary significantly, it is easier to use the protein extraction method due to the convenience of performing the MALDI-TOF–MS acquisitions at a later stage when all the isolates had been prepared. Several studies have confirmed that the protein extracts are stable when stored appropriately (Drevinek et al. 2012; Bohme et al. 2010a). The potential differences in media requirement for optimal pigment production were not considered and the standard medium YPS was used for all isolates as the focus was to obtain efficient screening. Some of the isolates did not produce good quality spectra and these were re-extracted from fresh cell material and also analysed using the direct smear method. However, most of these isolates still produced flatline spectra when the analyses were repeated. The protein spectra profiles of these isolates were then manually generated (omitting the quality criteria of the Biotyper) and analysed by SPECLUST (Alm et al. 2006). They could all be assigned to a few bacterial clusters indicating that specific bacterial taxa were more difficult to analyse than others and that the problems analysing them were not related to the experimental procedure as such (results not shown). One explanation of the poor spectrum quality is that the method and solvents used is not sufficient to lyse particular cell wall structures, as suggested to be the case with some Gram-positive *Bacillus* strains (Bizzini et al. 2011). One of the "problematic" clusters to analyse was an apparent agarocide phenotype. Another phenotype observation that might cause problems for the analysis is strains producing capsules or slime layers, structures surrounding the outside of the cell envelope. Problematic isolates are to be expected from environmental samples and it must be considered if the extra work load is expedient versus the gain of maintaining diversity during the study.

Fig. 2 Score oriented dendrogram calculated for MALDI-TOF MS spectra obtained for all included isolates generated by the default setting in MALDI Biotyper 2.0. The distance measure was set at correlation, and the linkage was set at complete. The default cut-off distance level at 500 is inserted. The identified strains are indicated by name. Species names followed by * were identified by 16S rRNA analysis from earlier studies and those followed by** were identified by both methods. **a** total dendrogram, **b** an enlarged cluster showing distinguished clusters of *E. amnigenus* and *S. plymuthica* with the common node below 500 (430) and again these two species have the distance to *Escherichia coli* slightly above 500. **c** example of well defined clusters with inter species differences all below 150 ▶

The bacterial isolates were clustered hierarchically based on the protein mass patterns and the resulting dendrogram is shown in Fig. 2a. Three replica MSPs of the external BTS standard, containing *Escherichia coli* DH5α were included to demonstrate the resolution of the method and that 100 % similar mass spectra are not achievable. However, as can be seen, the three technical replicates clustered with a distance below 50 on the relative 0–1,000 scale which is a satisfactorily result (see Fig. 2b). It was of interest to describe the bacterial collection in terms of diversity. For this purpose, a cut-off value had to be applied to be able to count the number of clusters. As mentioned in the Introduction, the cut-off value has only indicative properties at species level and not at the genus level. For example, *Enterobacter amnigenus* and *Serratia plymuthica* were separated at a distance level of ~470 while the intra-species variation is less than 100 (Fig. 2b). So by strictly utilizing a distance level of 500 for species differentiation, these two species would be counted as one. Nevertheless, it was decided to use a distance level of 500 to present a qualitative description of the bacterial diversity of the isolates picked randomly (the only criterion used was absence or presence of pigmentation) from the marine bacterial culture collection. The 373 isolates distributed into 55 clusters at the 500 distance level (Table 1A). Twenty-one clusters contained only one isolate (i.e. not clustering with any other isolate below distance level 500). In addition, 15 clusters were classified as low abundant taxa (less than 1 % of the total population but above 1 isolate, i.e. 2–4 isolates). This confirms the

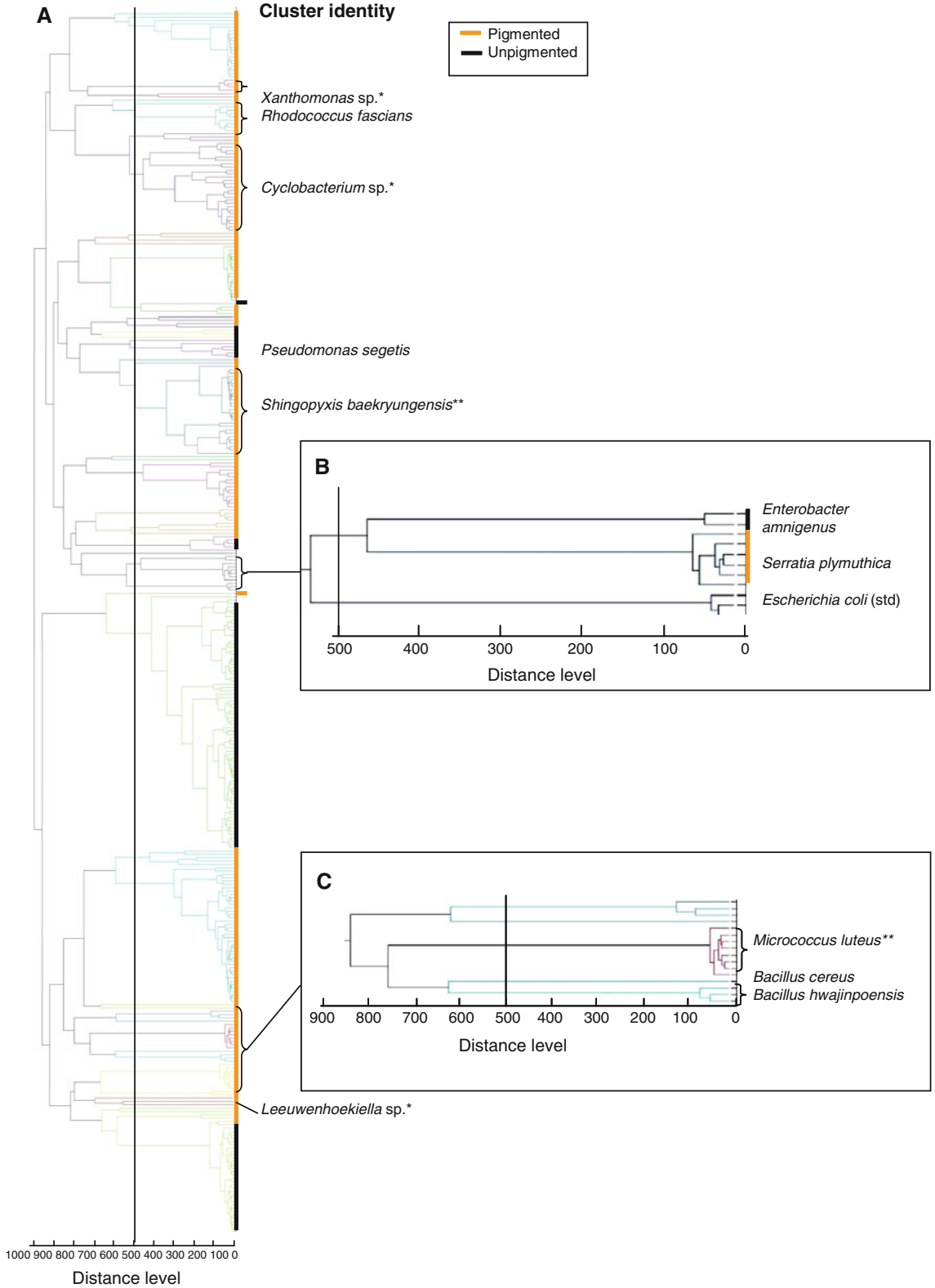

A
Cluster identity
Pigmented
Unpigmented
Xanthomonas sp.*
Rhodococcus fascians
Cyclobacterium sp.*
Pseudomonas segetis
Shingopyxis baekryungensis**
Leeuwenhoekiella sp.*
1000 900 800 700 600 500 400 300 200 100 0
Distance level
B
Enterobacter amnigenus
Serratia plymuthica
Escherichia coli (std)
500 400 300 200 100 0
Distance level
C
Micrococcus luteus**
Bacillus cereus
Bacillus hwajinpoensis
900 800 700 600 500 400 300 200 100 0
Distance level

Table 1 Overview of MALDI-TOF MS analysis divided in pigmented and pigmented bacterial groups (A), and bacterial isolates identified by Biotyper or 16S rRNA analysis (B)

A	Total	Pigmented (%)	Non-pigmented (%)	Mixed pigmentation (%)
Bacterial isolates	373	243 (65)	130 (35)	
Total clusters	55	45 (82)	9 (16)	1 (2)
Cluster with 2–4 isolates	15	13 (87)	2 (13)	
Cluster with 1 isolate	21	19 (91)	2 (9)	
B	Identified isolates (BioTyper Score Value[a])			
Pigmented	*B. cereus* (2.06)			
	M. luteus (1.94–2.20)			
	S. plymuthica (1.88–2.08)			
	Sphingopyxis baekryungensis (1.74–2.01)			
	R. fascians (1.73–1.89)			
	B. subtilis (1.81–1.84)			
	B. hwajinpoensis (1.76–1.93)			
	Cyclobacterium sp[b]			
	Xanthomonas sp[b]			
	Leeuwenhoekiella sp[b]			
Non-Pigmented	*E. amnigenus* (2.11–2.2)			
	Pseudomonas segetis (1.79–1.85)			

[a] The score of the microorganism identification is given as the lowest and highest value in the cluster

[b] Not in the Biotyper database, identified with 16S rRNA analysis in an earlier study (Stafsnes et al. 2010)

large diversity of bacterial isolates in the culture collection. 65 % of the 373 analysed isolates that generated good quality spectra were pigmented and distributed among 45 clusters (i.e. representing 82 % of the clusters). Furthermore, as the non-pigmented isolates constituted 35 % of the MALDI-TOF analysed samples but were distributed in only nine clusters (16 %) and the percentage further diminished when comparing low abundant clusters (13 %) and clusters with one isolate (9 %), the data shows that the pigmented bacteria selected have a greater diversity than non-pigmented bacteria. Only one cluster comprised both pigmented and non-pigmented isolates.

Since the Biotyper database contain mostly bacterial species of clinical interest, few of the marine isolates were, as expected, identified. The identified species are given in Table 1B together with the Biotyper score values. Their placements in Fig. 2a are also indicated. According to the criterion proposed by the manufacturer, a score value below 1.7 cannot allow reliable identification; a score value between 1.7 and 2.0 allows identification to the genus level ("secured to the genus"), and a score value higher than 2.0 is the set threshold for a match at the species level ("secured to the species"). Clearly, it would be useful if MALDI-TOF MS spectral data bases of environmental isolates become available, as is the case for clinical isolates. Alternatively, is would be desirable that a consortium of laboratories makes a joint effort to establish such a database, as the commercial potential is not as prominent as for clinical purposes.

Pigment analysis by LC-DAD-MS of selected isolates

Ninety-seven pigmented isolates were selected for pigment analysis by LC-DAD-MS. The selection was based on the MALDI-TOF MS analysis. Both single bacterium clusters and multiple bacteria clusters were represented, the latter by several isolates. A total of 82 unique pigments (i.e. unique absorption spectrum, m/z and retention time) were detected among the 97 analysed isolates. The number of pigments per isolate varied from 1 to 15, and most isolates produced between 3 and 5 pigments. Some representative pigment profiles, given as DAD chromatograms, are shown in Fig. 3. This clearly shows the large pigment diversity among the various isolates. Identification of

the pigments was outside the scope of this work but comparison with identified carotenoids from earlier work and purchased standards confirmed the ubiquitous presence of zeaxanthin, nostoxanthin and lycopene. Also the pigment prodigiosin was identified in extract from the isolate identified as *Serratia* sp., consistent with the well established production of this species by *Serratia* strains (Lewis and Corpe 1964).

The LC-DAD-MS data made the basis for a matrix of pigments versus isolates amenable for statistical analysis. Only the absence or presence of pigments was used in the definition of the pigment profile and not the relative production of each. Isomers and biosynthetic pathway intermediates (e.g. lycopene) were considered unique pigments. Figure 4b shows the pigment dendrogram obtained with the Unscrambler software using Average linkage clustering. The distance level is given as 0–10, were 0 is 100 % match. Twenty-eight different pigment profile clusters was identified by the software and manual evaluation. Most pigmented isolates belonging to the same cluster have a slight difference in the expression of pigment. For instance all the isolates identified as *Micrococcus luteus* have three common pigments and in addition one to three minor pigments (Fig. 3, profile f) resulting in a relative distance in the cluster up to 4 depicted in Fig. 4 as the purple cluster. The cluster depicted with a dotted line has a relative distance 7 and comprises four isolates which potentially show intra-species pigment variation (Fig. 3, profile a). These isolates were identified as *Rhodococcus fascians* and represent the most diverse pigment producing species in our study.

A new hierarchical clustering of MALDI-TOF MS spectra of the corresponding isolates, i.e. the 97 isolate sub-population, was performed and is shown in Fig. 4a. Thirty-one clusters were obtained at the

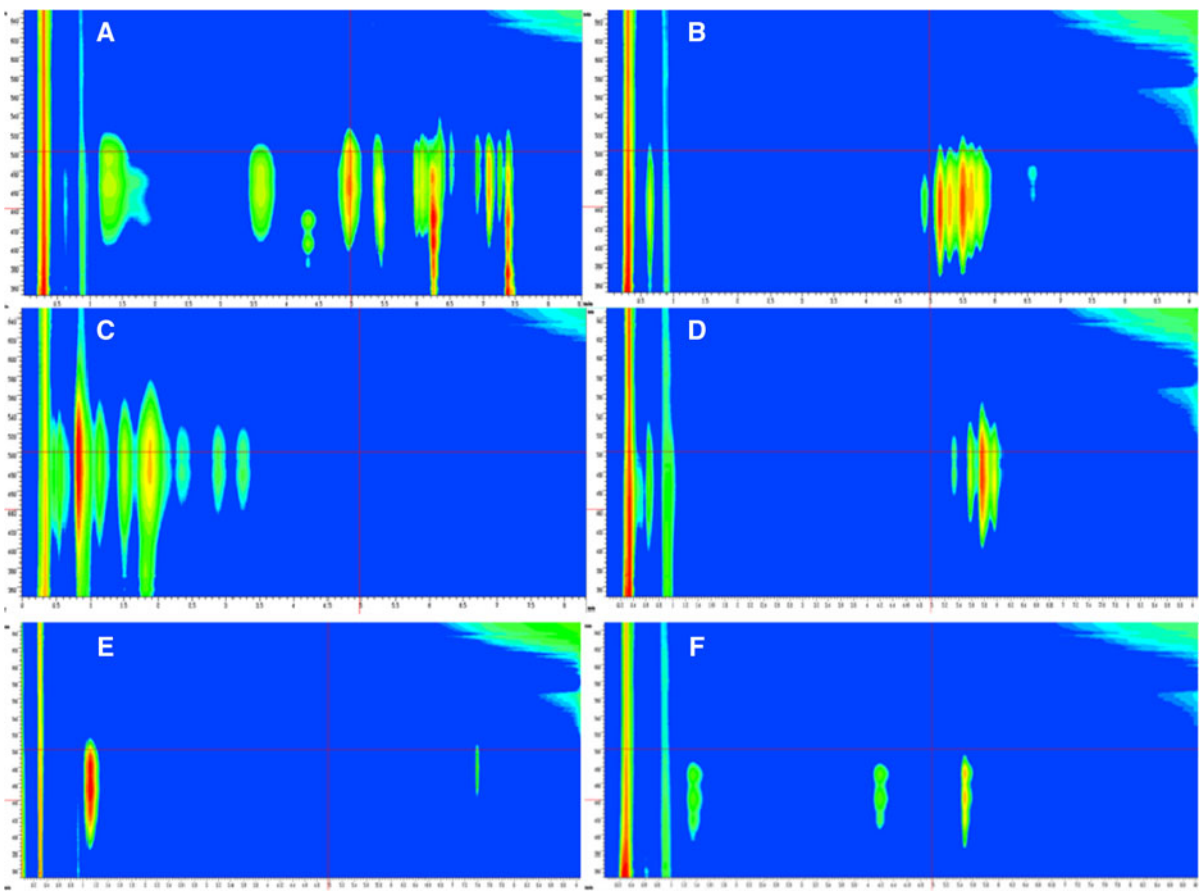
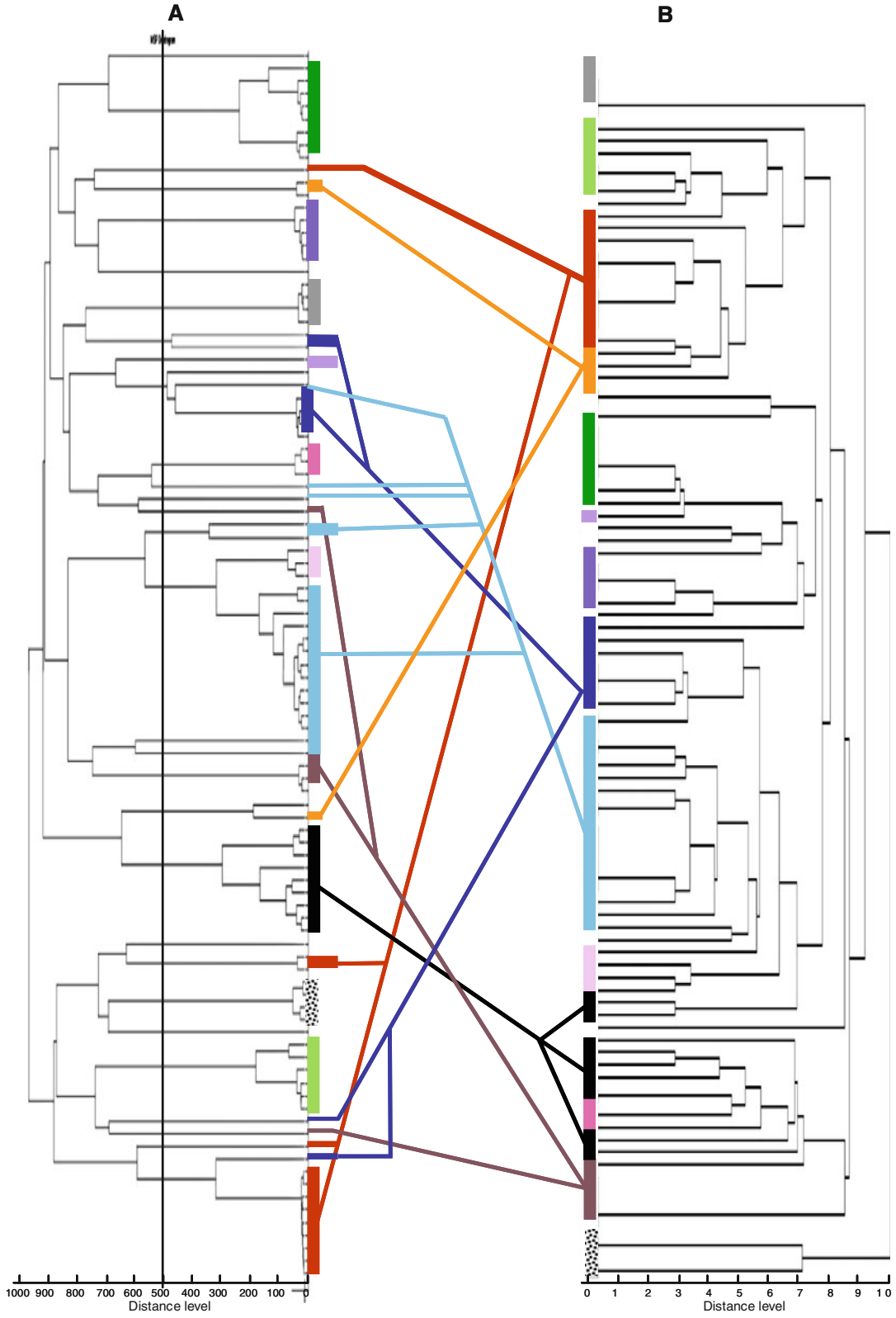


Fig. 3 Examples of pigment profiles with various numbers of pigments visualized with isoplots from LC-DAD analysis. **a** 14 pigments, *R. fascians*, **b** 8 pigments, *Xanthomonas sp.*, **c** 8 pigments, not identified, **d** 5 pigments, *B. hwajinpoensis*, **e** 2, not identified and **f** 3 pigments, *M. luteus*


A
B
1000 900 800 700 600 500 400 300 200 100 0
Distance level
0 1 2 3 4 5 6 7 8 9 10
Distance level

◀ **Fig. 4** The clustered pigment profiles (**b**) linked to the corresponding MALDI-TOF profile dendrogram (**a**). The distance level in the pigment cluster represents the difference in pigments, with 100–0 % match. Each pigment profile in one cluster is only identical (100 % match of all pigments) when they are connected at distance level 0, e.g. the cluster marked with a *dotted line* at the *bottom* contains four isolates, two and two grouped together with four identical pigments and three of these are identical for all four isolates resulting in distance level 7. When all isolates in one pigment profile corresponds to one MSP cluster in the MALDI dendrogram, no line is drawn, but is indicated by *color bars*. When the same pigment profile is distributed in more bacterial clusters *colored lines* are in addition drawn. When a bacterial cluster is only represented by one isolate and corresponds to a unique pigment profile, it is not indicated in the figure. (Color figure online)

distance level of 500. Eight of the pigment profiles have total correlation with bacterial clusters (when comprising more than one isolate). These are shown in Fig. 4 with identical colour-coded bars, and comprise 35 of the 97 isolates. In addition five clusters with only one isolate representing their unique pigment profile were detected. Furthermore, five pigment profiles can be found in more than one bacterial cluster. These have coloured lines (red, orange, dark and light blue and brown) linking isolate clusters (Fig. 4a) with respective pigment cluster (Fig. 4b). These pigment clusters all represent pigment profiles with only one or two pigments. As mentioned in the Introduction, most marine pigments are not unique for one bacterial species and the production of one or two pigments gives a weak basis for a bacterial fingerprint. It was therefore expected that this type of pigment profiles would be represented among several bacterial species. This is consistent with research from secondary metabolites in filamentous fungi where individual metabolites have been found in both phylogenetically closely related and distantly related species (Larsen et al. 2005). The pigment profile depicted with light blue comprised of 19 isolates and expressed only one pigment (Fig. 3e). In total 48 isolates represented in seven pigment clusters (25 %) are included in those isolates expressing only one or two pigments which does not have species-specific pigmentation.

In all the above mentioned clusters, no unique pigment profile would be lost if MALDI-TOF MS dereplication was applied to reduce the number of isolates for further pigment screening by LC–MS. MALDI-TOF MS can obviously not be used as dereplication tool for inter-cluster isolate elimination; however, it will still perform as an efficient pre-screening tool for intra-cluster elimination of replicas e.g. both one light blue and one red cluster in Fig. 4a contain over ten isolates each. Thus, for 90 % of the isolates (35 + 5 + 48) MALDI-TOF MS can be used as dereplication tool without losing pigment profile diversity. There is only one isolate cluster with several pigment profiles (black in Fig. 4). This cluster contains 9 isolates and the pigmentation profiles are distributed in three profiles at the relative distance of 5. Here, the MALDI-TOF MS would fail as a dereplication tool, as one would risk losing unique pigments by selecting just a few isolates for further LC-DAD-MS screening. But, this comprises less than 10 % of the population. Thus, the major conclusion from the comparison of MALDI-TOF MS-based isolate clustering with pigment clustering is that most pigment profiles follow species-specific classification when assuming that the resolution in the method lies at this taxonomic level.

Conclusion

MALDI-TOF MS is a powerful technique for characterizing large environmental bacterial collections, not only in context of characterization of microbial diversity but also as a dereplication tool to avoid superfluous work in exploitation of such culture collections. In particular, unique isolates are easily identified among a large number of related isolates. We showed that by using MALDI-TOF MS as a dereplication tool before pigment profiling of the isolates massive redundant analysis could be avoided. Only from one bacterial cluster, representing less than 10 % of the population, would we potentially have lost information by selecting one isolate from this particular cluster for further analysis as members of this cluster were shown to have different pigment profiles. In addition we showed that pigment profiling has potential to be used as a taxonomic marker at the species level when the isolates produce at least three different pigments. When fewer pigments are produced, no unique profile will be produced and therefore there is no value as a taxonomic marker since many different marine bacterial species are able to produce the same pigment.

Acknowledgments This study was financed by a grant from The Research Council of Norway and by the Norwegian

Defense Research Establishment. We thank Kjell D. Josefsen for advice on cultivation of the microbial culture collection.

References

AgnosTec (2012) SARAMIS^{TM} database for microbial identification. http://www.anagnostec.eu/products-services/reference-databases.html. Accessed 12.07 2012

Alm R, Johansson P, Hjernø K, Emanuelsson C, Ringnér M, Häkkinen J (2006) Detection and identification of protein isoforms using cluster analysis of MALDI-MS mass spectra. J Proteome Res 5:785–792

Andersen B, Hansen ME, Smedsgaard J (2005) Automated and unbiased image analyses as tools in phenotypic classification of small-spored *alternaria* spp. Phytopathology 95(9):1021–1029. doi:10.1094/phyto-95-1021

Appeltans W, Boucet P, Boxshall GA, Fauchald K, Gordon DP, Hoeksema BW, Poore GCB, van Soest RWM, Stöhr S, Walter TC, Costello MJ. (2012) World register of marine species. http://www.marinespecies.orgAccessed 26 Jan 2012

Bae SS, Kwon KK, Yang SH, Lee H-S, Kim S-J, Lee J-H (2007) *Flagellimonas eckloniae* gen. nov., sp. nov., a mesophilic marine bacterium of the family *Flavobacteriaceae*, isolated from the rhizosphere of *Ecklonia kurome*. Int J Syst Evol Microbiol 57(5):1050–1054. doi:10.1099/ijs.0.64565-0

Bizzini A, Jaton K, Romo D, Bille J, Prod'hom G, Greub G (2011) Matrix-assisted laser desorption ionization-time of flight mass spectrometry as an alternative to 16S rRNA gene sequencing for identification of difficult-to-identify bacterial strains. J Clin Microbiol 49(2):693–696

Bohme KI, Fernandez-No C, Barros-Velazquez J, Gallardo JM, Canas B, Calo-Mata P (2010a) Comparative analysis of protein extraction methods for the identification of seafood-borne pathogenic and spoilage bacteria by MALDI-TOF mass spectrometry. Anal Methods 2(12):1941–1947

Bohme KI, Fernandez-No C, Barros-Velazquez J, Gallardo JM, Calo-Mata P, Canas B (2010b) Species differentiation of seafood spoilage and pathogenic gram-negative bacteria by MALDI-TOF mass fingerprinting. J Proteome Res 9(6): 3169–3183

Daltonics (2012) MALDI Biotyper. www.bdal.com. Accessed July 12 2012

Dieckmann R, Graeber I, Kaesler I, Szewzyk U, von Dohren H (2005) Rapid screening and dereplication of bacterial isolates from marine sponges of the sula ridge by intact-cell-MALDI-TOF mass spectrometry (ICM-MS). Appl Microbiol Biotechnol 67(4):539–548

Dieckmann R, Helmuth R, Erhard M, Malorny B (2008) Rapid classification and identification of *Salmonellae* at the species and subspecies levels by whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry. Appl Environ Microbiol 74(24):7767–7778. doi: 10.1128/aem.01402-08

Drevinek M, Dresler J, Klimentova J, Pisa L, Hubalek M (2012) Evaluation of sample preparation methods for MALDI-TOF MS identification of highly dangerous bacteria. Lett Appl Microbiol 55(1):40–46. doi:10.1111/j.1472-765X.2012.03255.x

Dubois D, Leyssene D, Chacornac JP, Kostrzewa M, Schmit PO, Talon R, Bonnet R, Delmas J (2010) Identification of a variety of *Staphylococcus* species by matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol 48(3):941–945. doi:10.1128/jcm.00413-09

Dybwad M, Granum PE, Bruheim P, Blatny JM (2012) Characterization of airborne bacteria at an underground subway station. Appl Environ Microbiol 78(6):1917–1929. doi: 10.1128/aem.07212-11

Elssner T, Kostrzewa M, Maier T, Kruppa G (2011) Microorganism identification based on MALDI-TOF-MS fingerprints detection of biological agents for the prevention of bioterrorism. In: Banoub J (ed). NATO science for peace and security series A: chemistry and biology. Springer Netherlands, pp 99–113. doi:10.1007/978-90-481-9815-3_7

Fournier R, Wallet F, Grandbastien B, Dubreuil L, Courcol R, Neut C, Dessein R (2012) Chemical extraction versus direct smear for MALDI-TOF mass spectrometry identification of anaerobic bacteria. Anaerobe 18(3):294–297

Ghyselinck J, Van Hoorde K, Hoste B, Heylen K, De Vos P (2011) Evaluation of MALDI-TOF MS as a tool for high-throughput dereplication. J Microbiol Methods 86(3): 327–336. doi:10.1016/j.mimet.2011.06.004

Haigh J, Degun A, Eydmann M, Millar M, Wilks M (2011) Improved performance of bacterium and yeast identification by a commercial matrix-assisted laser desorption ionization–time of flight mass spectrometry system in the clinical microbiology laboratory. J Clin Microbiol 49(9): 3441. doi:10.1128/jcm.00576-11

Hansen ME, Carstensen JM (2004) Density-based retrieval from high-similarity image databases. Pattern Recogn 37(11): 2155–2164. doi:10.1016/j.patcog.2004.02.018

Hansen ME, Lund F, Carstensen JM (2003) Visual clone identification of *Penicillium* commune isolates. J Microbiol Methods 52(2):221–229

Hou Y, Braun DR, Michel CR, Klassen JL, Adnani N, Wyche TP, Bugni TS (2012) Microbial strain prioritization using metabolomics tools for the discovery of natural products. Anal Chem 84(10):4277–4283. doi:10.1021/ac202623g

Ichiki Y, Ishizawa N, Tamura H, Teramoto K, Sato H, Yoshikawa H (2008) Environmental distribution and novel high-throughput screening of APEO-degrading bacteria using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-MS). Journal of Pesticide Science 33(2):122–127

Jørgensen H, Fjærvik E, Hakvåg S, Bruheim P, Bredholt H, Klinkenberg G, Ellingsen TE, Zotchev SB (2009) Candicidin biosynthesis gene cluster is widely distributed among *Streptomyces* spp. isolated from the sediments and the neuston layer of the Trondheim Fjord, Norway. Appl Environ Microbiol 75(10):3296–3303. doi:10.1128/aem.02730-08

Klinkenberg G, Sletta H, Fjaervik E, Zahlsen K, Bruheim P (2011) Two-dimensional LC-MS fractioning and cross-matching of mass spectrometric data for rational identification of bioactive compounds in crude extracts. J Sep Sci 34(23):3359–3363. doi:10.1002/jssc.201100468

Koubek J, Uhlik O, Jecna K, Junkova P, Vrkoslavova J, Lipov J, Kurzawova V, Macek T, Mackova M (2012) Whole-cell MALDI-TOF: rapid screening method in environmental

microbiology. Int Biodeterior Biodegradation 69:82–86. doi:10.1016/j.ibiod.2011.12.007

Larsen TO, Smedsgaard J, Nielsen KF, Hansen ME, Frisvad JC (2005) Phenotypic taxonomy and metabolite profiling in microbial drug discovery. Natural Product Reports 22(6):672–695

Lewis SM, Corpe WA (1964) Prodigiosin-producing bacteria from marine sources. Appl Microbiol 12:13–17

Marklein G, Josten M, Klanke U, Müller E, Horré R, Maier T, Wenzel T, Kostrzewa M, Bierbaum G, Hoerauf A, Sahl H-G (2009) Matrix-assisted laser desorption ionization-time of flight mass spectrometry for fast and reliable identification of clinical yeast isolates. J Clin Microbiol 47(9):2912–2917. doi:10.1128/jcm.00389-09

Reichenbach H, Kohl W, Böttger-Vetter A, Achenbach H (1980) Flexirubin-type pigments in *Flavobacterium*. Arch Microbiol 126(3):291–293. doi:10.1007/bf00409934

Rettori D, Durán N (1998) Production, extraction and purificationof violacein: an antibiotic pigment produced by *Chromobacterium violaceum*. World J Microbiol Biotechnol 14(5):685–688. doi:10.1023/a:1008809504504

Rezzonico F, Vogel G, Duffy B, Tonolla M (2010) Application of whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry for rapid identification and clustering analysis of *pantoea* species. Appl Environ Microbiol 76(13):4497–4509

Sauer S, Freiwald A, Maier T, Kube M, Reinhardt R, Kostrzewa M, Geider K (2008) Classification and identification of bacteria by mass spectrometry and computational analysis. PLoS ONE 3(7):e2843. doi:10.1371/journal.pone.0002843

Simmons TL, Coates RC, Clark BR, Engene N, Gonzalez D, Esquenazi E, Dorrestein PC, Gerwick WH (2008) Biosynthetic origin of natural products isolated from marine microorganism–invertebrate assemblages. Proc Natl Acad Sci 105(12):4587–4594. doi:10.1073/pnas.0709851105

Stafsnes M, Josefsen K, Kildahl-Andersen G, Valla S, Ellingsen T, Bruheim P (2010) Isolation and characterization of marine pigmented bacteria from Norwegian coastal waters and screening for carotenoids with UVA-blue light absorbing properties. J Microbiol 48(1):16–23. doi:10.1007/s12275-009-0118-6

Takaichi S (2011) Carotenoids in algae: distributions, biosyntheses and functions. Mar Drugs 9(6):1101–1118

Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC, Bork P, Hugenholtz P, Rubin EM (2004) Comparative metagenomics of microbial communities. science 68(4):669–685

van Veen SQ, Claas ECJ, Kuijper EJ (2010) High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. J Clin Microbiol 48(3):900–907. doi:10.1128/jcm.02071-09

Veloo AC, Erhard M, Welker M, Welling GW, Degener JE (2011) Identification of gram-positive anaerobic *cocci* by MALDI-TOF mass spectrometry. Syst Appl Microbiol 34(1):58–62

Verroken A, Janssens M, Berhin C, Bogaerts P, Huang T-D, Wauters G, Glupczynski Y (2010) Evaluation of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of *Nocardia* Species. J Clin Microbiol 48(11):4015–4021. doi:10.1128/jcm.01234-10

Vynne NG, Mansson M, Nielsen KF, Gram L (2011) Bioactivity, chemical profiling, and 16S rRNA-based phylogeny of Pseudoalteromonas strains collected on a global research cruise. Mar Biotechnol 13(6):1062–1073

Welker M, Moore ERB (2011) Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 34(1):2–11. doi:10.1016/j.syapm.2010.11.013