In: Mass spectrometry for microbial proteomics
Eds: Haroun N. Shah, Saheer E. Gharbia, Vesela Encheva
John Wiley & Sons, London, UK *in press*

MALDI-TOF MS for Microbial Identification: Years of Experimental Development to an Established Protocol

Wibke Kallow[1], Marcel Erhard[1], Haroun N. Shah[2], Emmanuel Raptakis[3], Martin Welker[1§]

[1] AnagnosTec GmbH, Am Mühlenberg 11, 14476 Potsdam, Germany
[2] Health Protection Agency Centre for Infections, 61 Colindale Avenue, London NW9 5EQ, UK
[3] Shimadzu Biotech-Kratos Analytical, Trafford Wharf Road, Manchester M17 1GP, UK
[§] correspondence to m.welker@anagnostec.com

Abstract

Immediately after the first MALDI-TOF instrument became available, microbiologists have been investigating its potential as a platform for high throughput identification of microorganisms. A remarkable finding very early in these investigations was that despite the dynamic nature of the bacterial cell, components of the mass spectral profile were sufficiently stable and remained unchanged in spite of changes in environmental parameters. Further, mass spectral patterns have been found to be taxon specific, consequently, numerous methods were reported that purported to provide an alternative to current identification systems. The SARAMIS system described herein sequentially extracts core stable mass ions from analyses of multiple individual strains of a particular species to yield a 'SuperSpectrum', a list of biomarkers that are weighted according to their specificity from family to (sub)species levels. This approach has been used successfully to identify microorganisms from diverse phylogenetic lines of bacteria and fungi with considerable success. The protocol described has evolved of over years of experimental work to yield a robust system that can be readily applied for microbiological identification in a clinical diagnostic laboratory.

Identification of Microorganisms in Clinical Routine

A crucial step in the epidemiology and the successful therapy of any infectious disease is the identification of the causative microbe. For more than a century, clinical microbiology has relied on the isolation of the suspected pathogen from various samples such as stools, throat swabs, blood, or urine on selective growth media and an identification procedure that is based on the metabolic capacities of the isolate. An array of carbohydrate fermentation and enzymatic reactions are tested that generally involve a colour change of an indicator when a particular substrate is catabolised. The profile of positive and negative reactions is assumed to be characteristic for a bacterial taxon and is consequently used for identification. Modern microbial identification systems are miniaturized, combining some tens of reactions into a single strip or card to allow for high throughput analysis. The major shortcomings of these systems are the need to incubate isolates for several hours to obtain pure cultures, and a required pre-selection of tests. Although this is still the most commonly used method in clinical diagnostic laboratories, microbiologists have been seeking alternative methods for the identification of pathogens for decades.

A new era has dawned with the arrival of molecular methods such as the polymerase chain reaction (PCR) and nucleotide sequence analysis. In diagnostic and systematic microbiology today, analysis of genomic sequences is rapidly displacing biochemical tests for the provision of new characters for the circumscription of taxa. For example, a prerequisite for the description of a new species is the inclusion of the sequence of the 16S rRNA gene which now plays a pivotal role in microbial phylogeny. However, despite the widespread use of PCR and sequencing in all fields of microbiology, the technology is still lagging behind in clinical microbiology and is

largely restricted to research applications. On the other hand, the high sensitivity and specificity of molecular methods make them indispensable in modern microbiological laboratories, as for example in the detection of methicillin resistant *Staphylococcus aureus* (MRSA) by real-time PCR assays, or the identification of atypical or very rare pathogens.

Mass Spectrometry and Microbiology

The application of chemical analyses (referred to as chemotaxonomy) for the identification and classification of microorganisms has been explored extensively prior to molecular analysis. These were based on the characterisation of polar (eg. phospholipids) and non polar lipids such as respiratory quinones (eg. ubiquinones and menaquinones) and long-chained cellular fatty acids [1]. The structure of these lipids were challenging and ushered in a period of intense mass spectral analysis to characterise the vast array of lipids present in the microbial kingdom. While these methods provided characters at the genus and species levels, pyrrolysis mass spectrometry was introduced as a means of typing bacterial isolates [2]. Such approaches were motivated by the need for a rapid method to identify pathogens in only a fraction of the time required for biochemical tests. However, because of the limitations of the technology at that time, mass spectral approaches were confined to the detection of organic molecules in a mass range up to 1,500 Da [3]. Detection of larger molecules was hitherto only possible with techniques such as plasma desorption mass spectrometry [4]. But this was about to change dramatically within a few years with the invention of Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry (MALDI-TOF MS) and its success in many fields of life sciences has been phenomenal.

The laser desorption and mass spectral analysis of large biomolecules was developed simultaneously by two research groups in Japan and Germany. While the group of Tanaka could successfully detect proteins up to m/z 100,000 Da by direct laser ionization [5], Karas and Hillenkamp [6] relied on a light absorbing matrix, and achieved similar results by this method. A matrix effect on the desorption rate was observed earlier for smaller molecules [7] and further studies quickly directed the search for matrix candidates to a few small organic molecules that are still the primarily used ones in MALDI-TOF MS applications today, *viz.* cinnamic acid derivatives [8] and 2,5-dihydro benzoic acid [9]. Accuracy and resolution could be significantly improved by the introduction of delayed ion extraction [10], which compensates for the variability in initial ion velocity [11]. The sensitivity of MALDI-TOF MS for the detection of large proteins was rapidly increased to the femtomolar range [12] and further, aided by sophisticated handling procedures to the zeptomolar range [13]. Within less then a decade, MALDI-TOF MS developed into a widely applied methodology in diverse fields of life sciences [14], promoted by a number of major advantages it had compared to other mass spectral technologies. These included the possibility to detect unfragmented large molecules, the speed with which a full scan over a wide mass range can be achieved, and the simplicity of sample preparation. For the analysis of whole cells and crude extracts, an outstanding advantage of MALDI compared to other ionization techniques such as electrospray ionization (ESI) or fast atom bombardment (FAB), is the fact that in MALDI-TOF MS predominantly singly charged ions are detected [15] which simplified the interpretation and treatment of mass spectral data considerably [16].

Mass Spectral 'Fingerprints' of Whole Cells

The possibility of introducing whole cells into a mass spectrometer and detecting biomolecules in a mass range extending to several kilodaltons was immediately recognised by microbiologists. Shortly after the first MALDI-TOF MS was commercially available (by Vestec and Kratos) the first reports on intact cell mass spectrometry of bacteria were published, some of which highlighted its potential for microbial diagnostics [17-21]. Essentially these studies showed that mass spectra of whole cells of bacterial strains revealed patterns of mass signals that were reproducible and specific for strains or species. Because cells could be analysed after minimal processing and preparation, required only minute biomass, either as a cell suspension or cells placed directly on a target plate, the implications for diagnostic microbiology were immediately evident and numerous studies ensued.

The mass ranges that were selected in early studies varied from *m/z* 500-2200 [18] to *m/z* 2000-20000 [22]. In the lower mass range, constituents of the cell wall are detected, and also, but to a lesser degree protein components [23]. The lower mass range was also used for the typing of potentially toxic cyanobacteria [24] and proved to be suitable for the metabolic typing of sub-specific taxonomic units in natural populations [25].

The possibility that mass fingerprints could be used to classify and eventually identify

unknown microbial isolates was pursued in the following years by a number of groups. Most studies focused on the detection of proteins in mass ranges spanning from above *m/z* 2000 to below *m/z* 25,000. The identity of proteins detected at this time was not very clear because only a few genomic sequences of microorganisms were available. It was generally assumed that the proteins desorbed from whole cells, that were not subjected to typical mechanical or chemical lysis, were attached to the cell surface [23]. However, studies on membrane-associated proteins of *E. coli* K12, revealed that the molecular mass of most of these proteins exceeded 20 kDa [26] and were likely to be intracellular proteins.

Comparative MALDI-TOF mass spectral studies of isolated ribosomal subunits with those of whole cell mass spectra of *E. coli* K12 revealed that they shared a large number of mass signals that are commensurate with ribosomal proteins [27] (Figure 1). This large number of ionised ribosomal proteins is in accord with the high level of these proteins (*ca.* 30% of total proteins) of a cell in its exponential growth phase. Furthermore, studies on isolated ribosomes also revealed that many proteins were modified post-translationally, and that the observed mass signals were in agreement with the general rules for methionine cleavage [28]. Such studies emphasised that direct 'translation' of genomic data in mass patterns is not straightforward.

Further microbial taxa have been studied in detail to reveal the identity of proteins detected by whole cell mass spectrometry. A high level of observed peaks in mass spectra could be assigned to ribosomal proteins in *Helicobacter pylori* [29], *Saccharomyces cerevisiae* (together with mitochondrion related proteins) [30], *Campylobacter jejuni* [31], *Lactobacillus plantarum* [32], and *Pseudomonas putida* [33]. With the growing number of completed genomes in databases, the assignment of mass signals to proteins will be extended to further species. In Figure 2, a mass spectrum of a clinical isolate of *Streptococcus pneumoniae* is shown with the mass signals matching to theoretical masses of ribosomal proteins obtained from the SwissProt database. Other proteins identified in whole cell mass spectra include for example, DNA-binding proteins [31] and a protein in *E. coli* that correlates with the possession of an F-plasmid (*m/z* 9743) [34]. The latter example also underlines the limits of *in silico* studies, since no candidate among some thirty proteins encoded on the F-plasmid had a sequence matching the observed mass signal when considering only modifications such as methionine cleavage or methylation. The *in vivo* protein could be further modified for example by cyclisation or formation of di-sulfide bridges, or cleaved from a larger premature protein.

The detection of ribosomal and other small structural proteins (i.e., proteins without a direct metabolic function) in whole cell MALDI-TOF mass spectra of various microbial taxa is very promising for the development of identification systems for clinical applications. To this end, two major objectives need to be addressed. Firstly, the reproducibility of mass fingerprints acquired under various culture conditions, on different instruments, or from multiple strains of a single species and, secondly, the discriminating power of mass fingerprints among recognized species.

Reproducibility of Mass Spectral Fingerprints

A principle problem for the development of a standard protocol using mass spectrometry is the fact that even replicate analyses of a single clone do not result in identical mass fingerprints. This applies to the individual peak's height as well as to the presence and absence of peaks, primarily of those with low intensities. The reproducibility of mass spectral fingerprints therefore has been the subject of a number of studies assessing biological, physical, and chemical effects on whole cell mass spectra. Wang *et al.* [35] demonstrated that variations in the sample preparation procedures and in the experimental conditions resulted in a significant change in the observed spectra, especially in relative peak intensities. Thus, a number of peaks were observed in the spectra from the same bacterial sample obtained under different conditions, especially when acidic extraction was employed. Similar observations were made by Williams *et al.* [36] who also reported that acidification of the matrix solution with 1-2.5 % TFA yielded the best results in terms of peak number and reproducibility with different matrices. The shape of recorded peaks is itself dependent on a number of factors such as sample concentration and laser intensity, and consequently these factors need to be optimised for a given instrument [37].

Since peak recognition partly depends on the software settings rather than on the hardware components, a sample preparation procedure that yields reproducible results on one MS is likely to be suitable for others. This was assessed in an inter-laboratory study where a single sample was analyzed in three laboratories on different mass spectrometers [38]. Again, a number of mass sig-

nals were invariably recorded in all fingerprint spectra, while others were recorded with differing frequency in different laboratories. Thus, to a certain degree differences in whole cell mass spectra of a single bacterial sample can be attributed to physical and chemical factors. This is important to note since spectrum-to-spectrum variability could be mistakingly taken to indicate biological differences.

Biological variability is expected in whole cell mass spectral fingerprints due to the dynamic regulation of protein expression in response to growth conditions and culture age. Thus, for *E. coli* it was shown that mass spectral profiles changed with cultivation time [39]. Assuming that the mass ions of the ribosomal proteins are recorded in mass spectral fingerprints of many species, it is expected that the signal intensity of these proteins will be highest in the exponential growth phase. Figure 3 shows the mass spectra of whole cells of *Proteus mirabilis* recorded after cultivation times varying between 16 to 120 hours. Although the quality of the mass fingerprint declines considerably after 48 hours (in terms of peak number and resolution) the major mass signals were still recorded and allowed automated identification (see below). However, our own observations have shown that the effect of culture conditions on mass spectral fingerprints is less pronounced than would be expected considering the complex system of protein expression in living cells. This was confirmed by Valentine *et al.* [40] who examined the mass spectra of several species using different culture media and Wunschel *et.al* who additionally analysed cells grown at different temperatures, pH, and growth rates [41].

Species and Strain Discrimination by Mass Spectrometry

Early studies that explored the use of mass spectral fingerprints tested the discriminating power on rather distantly related taxa, however, as confidence grew in the potential of the technique, more closely related taxa were studied (Table 1). *Bacillus* spores and cells have been studied intensively because of their potential use as bioterrorism agents and the need for a fast, safe and reproducible method to identify them in suspicious samples. It proved that with MALDI-TOF MS (and other mass spectral techniques) cells and spores of one of the most dangerous species, *B. anthracis* and the generally inseparable species, *B. cereus*, could be differentiated [42-44]. Reliable biomarkers were reported as 'small acid soluble proteins' (SASP) that were characteristic of a number of strains [45] and had the potential for identification of cells and spores [46,47].

Lynn *et al.* [48] studied a number of Enterobacteriaceae of different genera and showed that a small number of mass signals were recorded for all strains, e.g., m/z 4364 (ribosomal protein L36), 5380 (ribosomal protein L34), and 6856 (YcaR). For particular genera, species, and strains-specific mass signals could be identified that allowed the discrimination of *E. coli*, *Klebsiella pneumoniae*, *Providencia rettgeri*, and *Salmonella typhimurium*. Such studies illustrated that mass spectral fingerprints can potentially be used for the classification of bacterial strains. Further studies showed that environmental *Escherichia coli* isolates could be classified by MALDI-TOF MS according to their origin. Thus, avian *E. coli* isolates could be distinguished from those of bovine origin [49].

In other studies, clinically relevant *Streptococcus* species were analysed by MALDI-TOF MS. Identification and differentiation of *Streptococcus* species is generally challenging by routine clinical methods while some groups of species sharing more than 99% nucleotide sequence similarity in the 16S rRNA gene. However, using MALDI-TOF MS mutans streptococci could be discriminated down to the subspecies level [50] permitting the reclassification of some strains. Similarly, *Streptococcus pyogenes* strains could be identified correctly and differentiated into invasive and non-invasive isolates [51]. *Streptococcus* strains belonging to the viridans group were analysed by Friedrichs *et al.* [52] and shown that closely related species such as *Str. mitis* and *Str. oralis* could be reproducibly separated.

Some twenty clinically relevant species and subspecies of *Staphylococcus* were studied by Carbonelle *et al.* [53] who reported that each species had a distinct mass spectral fingerprint that allowed its discrimination, allowing the high throughput identification of *Staph. aureus*[54]. Discrimination between methicillin-resistant (MRSA) and methicillin-susceptible (MSSA) *Staphylococcus aureus* strains by MALDI-TOF MS was reported by Edwards-Jones *et al.* [55] and Du *et al.* [56]. Both types of strains differed in mass signals around m/z 2500 and their assignment was in accordance with the results of a *mecA*-specific PCR assay. A few exceptions, however, were observed and the discrimination of MRSA and MSSA by MALDI-TOF MS could not be confirmed by the analysis of further strains (AnagnosTec, unpublished data; [54,57]). Nonetheless, considering the high costs of MRSA identification by real-time PCR assays, an alternative

methodology would be advantageous. Since MRSA are not a monophyletic unit but have arisen from different MSSA lineages independently [58] new approaches will require a different strategy. Direct detection from whole cells of gene products of the staphylococcal chromosomal cassette (SCC*mec*) responsible for the β-lactam resistance is, however, unlikely to be feasible. The penicillin binding protein 2a, for example, has a mass of 78 kDa, far beyond the mass range generally used for mass fingerprinting. Methods involving surface enhanced laser desorption/ionisation (SELDI TOF MS) shows promise [54]. This technique has also been used to delineate subspecies of *Francisella tularensis*, the causative organism of tularaemia [59].

A number of *Mycobacterium* species were found to be distinguishable by mass spectrometry, and strains could even be discriminated when mass signal intensity was accounted for in addition to simple presence/absence data [60].

As agents of bioterrorism usually do not emerge in laboratory settings, a procedure is required to make any potentially hazardous substance innocuous immediately, without disabling its biomarkers for subsequent identification. A number of *Bacillus* (spores and cells) and *Yersinia* species were studied for the applicability of a thorough inactivation protocol prior to mass spectral characterization [61]. Cells and spores were treated with 80% trifluoroacetic acid resulting in complete inactivation that did not significantly deteriorate the quality of mass spectral fingerprints, which could still be used for identification.

Compared to bacteria, far less attention has been given to the mass spectral analysis of fungi. However, the few studies undertaken so far indicate that pathogenic fungi can also be identified by whole cell MALDI-TOF MS. Closely related wood decay fungi could be discriminated and identified [62]. Clinical isolates of dematophytes could be automatically identified by matching the isolates' spectra to a spectral database [63]. An infection with *Fusarium proliferatum* was confirmed by mass spectral typing and comparison to reference strains [64]. A major obstacle in the mass spectral identification of fungi is the variability in a strain's mass fingerprint due to cell differentiation and spore formation.

Among the protists, some notorious pathogens are relevant for public health surveillance such as *Giardia lamblia* or Cryptosporidia. Both taxa have been analysed by MALDI-TOF MS and the results suggest that mass spectral fingerprints can also be applied for their correct identification [65,66].

Proof of principle has been demonstrated for several species (see Table 1) but the shortcomings of some studies prevent their direct transfer to a diagnostic laboratory. Firstly, in many studies only a limited number of strains were tested, which is unlikely to fully represent the intra-specific diversity of a given taxon, and secondly, because these identification principles have only rarely been tested for unknown samples

For mass fingerprints to be used for the routine identification of unknown samples, methods have to be developed that treat the mass spectral information in a reproducible and standardized manner, which can then be compared with reference mass spectra of well characterized strains.

Pattern Matching Approaches for automated Identification

In general, the comparison of mass spectral fingerprints is done by mathematical approaches based on various algorithms for cluster analyses, or multivariate approaches such as principal component analysis. The key issue for a successful method is not the particular algorithm that is applied to the processed data since different algorithms may lead to very similar results [67]. However, the processing of the raw data and the identification of mass signals is of paramount impotance.

Different approaches have been proposed for pattern matching procedures. Arnold and Reilly [68] applied a cross-correlation analysis to mass spectra of 25 *E. coli* strains by dividing the full spectrum in a varying number of intervals. By this procedure, individual strains could be discriminated even when the corresponding mass spectra appeared very similar. Demirev *et al.* [69] proposed a combination of mass spectrometry and protein database search to assign signals in the mass spectrum of an unknown bacterial sample to a protein sequence translated from internet-accessible genomic sequences. Jarman *et al.* [70] developed an algorithm for the extraction of fingerprints from mass spectra (*m/z* 1,000-10,000) of bacteria considering the variability in intensity, frequency, and accuracy of mass signals in replicate analyses of an individual strain. By doing so, a fingerprint was calculated that accounts for differences in replicate mass spectra caused by physiological regulation, analytical error and detection limit. The extracted fingerprint data was applied to the comparison of blind samples with library reference fingerprints of five species by calculating similarities between an unknown and a reference fingerprint based on the absence

and presence of individual biomarkers [71]. For the blind samples, the same strains and mixtures thereof with *Shewanella alga* cells were used, and fingerprints calculated from repeated analyses. For all species except *Bacillus cereus*, a correct identification was achieved when these were part of the mixture, and for all strains correct negative results were returned in the case of their absence. Bright *et al.* [72] applied a pattern recognition algorithm to mass spectra in a range of *m/z* 500-10,000. The software tool MUSE translated each quality-controlled mass spectrum into a single point vector in an n-dimensional space. The spectra of replicate samples of a strain had a location in an n-dimensional space close to, but not identical to each other. As a reference library the mass spectral data of 35 strains from 20 species were included. When the same strains were analysed as blind samples, the correct match was achieved for more than 95% on the species level. A hierarchical clustering algorithm was applied by Hsieh *et al.* [73] in combination with analysis of variance (ANOVA) to extract biomarkers from multiple isolates belonging to six human pathogens. The set of specific markers was then reduced to 2-4, which was sufficient to identify isolates of the same species correctly.

Mass Spectral Identification of Microorganism – Requirements for Routine Diagnostics

The proof of principle of mass spectral identification of microorganisms has been shown for various taxa (see Table 1). This is an essential prerequisite for diagnostics in clinical microbiology but for daily routine use there are further demands. Thus, mass spectral identification needs to be

- simple and fast in handling
- robust to account for variations and variability in culture conditions
- reproducible to allow identifications at different locations
- applicable to the majority of clinically relevant microorganisms
- economic to allow identifications at competitive costs

The measuring process itself, as described in most publications, is straightforward. Several extraction solvents and matrix compounds have been tested for their effect on resulting mass fingerprints [74,75]. Despite various suggestions for extraction procedures, cells of most taxa can generally be applied directly to sample targets, either from solid media or from suspensions. One particular exception, yeasts, have been shown to result in considerably improved mass fingerprints after an additional extraction step with formic acid [30], but for most bacteria, extraction performed on the sample spot by adding matrix solution can be considered as sufficient. Extraction procedures involving more steps may lead to increased reproducibility, but at the expense of a significant increase in handling time.

Once raw mass spectra have been obtained, mass signals need to be processed by peak detection algorithms to extract the relevant information for further analysis. Although this is performed by all software packages supplied with a mass spectrometer, differences in peak recognition algorithms can have a considerable effect on the processing and need to be adjusted with care [76]. Raw data could be used for spectral comparison directly, e.g., by the BioNumerics software package as it has been applied to mass spectra of *Burkholderia cepacia* complex species by Vanlaere *et al.* [67]. A major disadvantage of this approach is the extended computing time for large datasets. For practical reasons therefore, the extraction of peak data from raw mass spectra substantially facilitates the downstream data handling, and can be generally performed with the MS software with appropriate settings for baseline correction, noise filtering, etc.

Finally, the identifying step involves the comparison of the mass spectral fingerprint of an unknown sample with the mass spectral fingerprints of well characterized strains. The principles applied to the identification of microorganisms based on mass spectral fingerprints are essentially the same as for similar problems where a pattern recognition is required, for example, restriction fragment length polymorphism analysis. A major obstacle in this process is the fact that every mass signal in a spectrum has an analytical error. This is in contrast to the comparison of (nucleotide) sequence data where only discrete values are compared. In MALDI-TOF MS several factors contribute to the analytical error, e.g., small variability of acceleration voltage, shape of matrix crystals, and peak recognition by the processing software. For most linear MALDI-TOF mass spectrometers an error of some 500 ppm is considered as acceptable, i.e., a deviation of 5 Da for *m/z* 10,000 . A higher precision can be achieved by internal standards but for ICMS this is not applicable because the calibrants potentially interfere with biomarkers. In consequence, a particular biomarker is detected with a variability in *m/z* values that has to be taken into

consideration when pattern matching algorithms are applied.

The most important part of a system for automated mass spectral identification of microorganisms is a validated reference database [57,77,78]. The reference database should contain not only the mass spectral data of pathogens but also of related, non-pathogenic species that are frequently found in a clinical environment to avoid the danger of false positive identifications that could lead to a superfluous and potentially stressful therapy. From these considerations, it is evident that a database needs to contain reference spectra of hundreds of microbial species.

Such a database was established by AnagnosTec in 1998 and is constantly amended for rare and newly emerging pathogens, microbes relevant in veterinary medicine, and environmental species. In the next section the principle and workflow for automated mass spectral identification of clinical microorganisms using SARAMIS (Spectral archiving and microbial identification system) is briefly explained and examples of the application in routine laboratories are given.

Automated Mass Spectral Analysis of Microorganisms in Clinical Routine Diagnostics

The identification system SARAMIS consists of a database containing reference spectra, and software that allows the comparison of the mass fingerprint of an unknown sample to the reference spectra and was developed by AnagnosTec[79]. The same principle is applied in the BioTyper software package developed by Bruker Daltronics [80].

As reference spectra, SARAMIS uses so-called "SuperSpectra" that have been calculated from replicate mass spectra of multiple isolates of a single species or sub-species. To generate SuperSpectra, high quality mass spectra of a species of interest are collected from multiple strains grown under varying conditions and in different laboratories. Raw mass spectra are then processed to yield peak lists, that are subjected to cluster analysis by applying a single link agglomerative clustering algorithm that allows a pre-set analytical error, generally 500 ppm. Each cluster represents a mass fingerprint type of a species as shown in Figure 4. When numerous strains of a single species have been analysed, cluster analysis may reveal sub-specific units and in that case, multiple SuperSpectra can be calculated. In the next step, a consensus spectrum is calculated containing only these mass signals that have been recorded in a frequency of 50-100% to exclude unspecific signals. This threshold frequency needs to be taxon-specific with a dataset at hand that allows estimation of the intraspecific variability. Consensus spectra inevitably contain peaks that are specific for higher taxonomic units such as genera or families which can give mass spectral fingerprints of members of respective taxa a high background similarity (see, for example, Enterobacteriaceae [48], Figure.4). For species identification respective mass signals cannot be reclaimed and are given a low weighting. For the calculation of SARAMIS SuperSpectra such non-specific mass signals are identified by comparing each *m/z*-value to *m/z*-values in the entire database of reference spectra (> 100,000; 2009). Before SuperSpectra are released for automated identification, each mass signal contained therein is weighted according to its specificity for the taxon it represents, e.g., a mass signal that has been recorded exclusively with high frequency in one species receives a high value while one that has been recorded in multiple species of the same family receives a lower value.

When a mass spectral fingerprint is matched to the SuperSpectra in the database, weighting values for each matching mass signal are summed up and as a result those SuperSpectra for which the highest specific concordance was established are listed. The sum of the mass signal specific match values can then be translated into a confidence value for identification.

The performance of SARAMIS was evaluated in several clinical diagnostic laboratories [81], and the results of one study presented here. In 2007, routine identification of isolates from human urine samples was performed with two systems in parallel in the laboratory of Dr. Stein (Mönchengladbach, Germany). Samples were analysed by the VITEK 2 system already established in the laboratory as well as by MALDI-TOF MS/SARAMIS that was set up in the laboratory in 2007. A total of 569 bacterial isolates were analysed by both methods representing the commonly encountered, clinically relevant taxa, including diverse species of Enterobacteriaceae, gram-positive cocci, and nonfermentative species. Both methods yielded consistent results for 96.6% of all isolates, including 6.3% for which the SARAMIS identification was correct but only with a confidence level of 80-90% (Table 2). The best hit rates were achieved for Enterobacteriaceae and Nonfermenters (98 and 100%, respectively) while for staphylococci and Entero/streptococci only 92.4 and 82.8% correct identifications were attained, respectively. For these two groups, the highest number of unidentified isolates or isolates for which automatic data

acquisition did not yield a mass fingerprint were encountered. Thus, although the mass spectrum had passed the quality control, when a match to SuperSpectra accounts to < 70% confidence, SARAMIS issues a non identification result. The main reason for a negative result is generally gaps in the database, either encompassing rare species for which sufficient date have not yet been collected or subtypes of well-characterized species for which the mass fingerprint deviates considerably from typical mass fingerprints of that species. For example, in *Streptococcus* sp. the transition between species is frequently accompanied by new mass fingerprint types (own, unpublished data). For 1.1% of the isolates, no mass fingerprint could be generated by automated spectrum acquisition, but in the manual data acquisition mode for part of these samples a mass spectrum could be obtained. One of the most important outcomes of the evaluation was that not a single sample was mis-identified. These results clearly underline that a mass-spectrometry based identification system is now readily available as a diagnostic tool for the clinical laboratory and can be implemented after a short familiarisation phase for the routine identification of clinical samples.

Nonetheless, it has to be emphasized that mass spectrometry based identification of microorganisms has its limits. This is partly due to the fact that bacterial identification is based indirectly on a genomospecies concept (ie. genes translated into proteins that were detected) while classical clinical microbiology is based on phenotypic characters derived from metabolic profiles [82]. A typical example is the *Shigella* species that are recognized by taxonomists as polyphyletic lineages within the species *Escherichia coli* [83]. The main phenotypic characteristics of *Shigella* sp. are that they are non-motile and cannot metabolise lactose; characteristics that also occur, however, in enteroinvasive *E. coli* strains [84] and which cannot be determined directly by intact cell mass spectrometry. On the other hand, species discrimination by mass fingerprinting can be straightforward for taxa that are poorly discriminated by biochemical methods [85], for example for species of the *Burkholderia cepacia* complex [67]. Consequently, the prospects and capacities of MALDI-TOF MS based and biochemical microbial identification procedures are not completely congruent and the application of the former in routine diagnostics will require some training by the medical staff to interpret the data. However, traditional biochemical tests are now giving way to genomics approaches for species identification especially since the introduction of 16S rRNA gene sequencing. Throughout the entire microbial kingdom, many groups of biochemically similar species have been further resolved and clarified based on 16S rRNA sequences and the current edition of Bergey's manual has adopted this approach for a 'road map' of the microbial kingdom. There is therefore a great opportunity to interphase a MALDI-TOF MS based method with 16S rRNA sequencing to provide a modern and reliable approach to microbial identification. The simplicity, speed, and reliability of mass-based identification techniques are indisputable advantages, and it is expected that the methodology will be strengthened by the large number of studies in progress and in the future will become part of the work of routine diagnostics laboratories.

Acknowledgements

We whish to thank Dr. Lothar Kruska for making the data of the SARAMIS evaluation available.

Cited references

1. J.P.Anhalt and C.Fenselau, Identification of bacteria using mass spectrometry, *Anal. Chem.* **47**, 219-225 (1975).
2. H.R.Schulten, H.D.Beckey, H.L.C.Meuzelaar and A.J.H.Boerboom, High resolution field ionization mass spectrometry of bacterial pyrolysis products, *Anal. Chem.* **45**, 191-195 (1973).
3. D.N.Heller, R.J.Cotter, C.Fenselau and O.M.Uy, Profiling of bacteria by Fast Atom Bombardment Mass Spectrometry, *Anal. Chem.* **59**, 2806-2809 (1987).
4. B.Sundqvist, A.Hedin, P.Hakansson, I.Kamensky, M.Salehpour and G.Säwe, Plasma desorption mass spectrometry (PDMS): limitations and possibilities, *Int. J. Mass Spectrom. Ion Processes* **65**, 69-89 (1985).
5. K.Tanaka, H.Waki, Y.Ido, S.Akita, Y.Yoshida and T.Yoshida, Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry, *Rapid Comm. Mass Spectrom.* **2**, 151-153 (1988).
6. M.Karas and F.Hillenkamp, Laser desorption ionization of proteins with masses exceeding 10,000 Da, *Anal. Chem.* **60**, 2299-2303 (1988).
7. M.Karas, D.Bachmann, U.Bahr and F.Hillenkamp, Matrix-assisted ultraviolet laser desorption of non-volatile compounds, *Int. J. Mass Spectrom. Ion Processes* **78**, 53-68 (1987).
8. R.C.Beavis and B.T.Chait, Cinnamic acid derivatives as matzrices for ultraviolet laser desorption mass spectrometry, *Rapid Comm. Mass Spectrom.* **3**, 432-435 (1989).
9. K.Strupat, M.Karas and F.Hillenkamp, 2,5-dihydroxybenzoic acid: a new matrix for laser desorption-ionization mass spectrometry, *Int. J. Mass Spectrom. Ion Processes* **111**, 89-102 (1991).
10. U.Bahr, J.Stahl-Zeng, E.Gleitsmann and M.Karas, Delayed extraction time-of-flight MALDI mass

spectrometry of proteins above 25000 Da, *J. Mass Spectrom.* **32**, 1111-1116 (1997).
11. M.Glückmann and M.Karas, The initial ion velocity and its dependence on matrix, analyte and preparation method in ultraviolet matrix-assisted laser desorption/ionization, *J. Mass Spectrom.* **34**, 467-477 (1999).
12. M.Karas, A.Ingendoh, U.Bahr and F.Hillenkamp, Ultraviolet-laser desorption/ionization mass spectrometry of femtomolar amounts of large proteins, *Biol. Mass Spectrom.* **18**, 841-843 (1989).
13. S.Jespersen, W.M.A.Niessen, U.R.Tjaden, J.van der Greef, E.Litborn, U.Lindberg, J.Roeraade and F.Hillenkamp, Attomole detection of proteins by matrix-assisted laser desorption/ionization mass spectrometry with the use of picolitre vials, *Rapid Comm. Mass Spectrom.* **8**, 581-584 (1994).
14. F.Hillenkamp and M.Karas, Matrix-assisted laser desorption/ionization, an experience, *Int. J. Mass Spectrom.* **200**, 71-77 (2000).
15. M.Karas, M.Glückmann and J.Schäfer, Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors, *J. Mass Spectrom.* **35**, 1-12 (2000).
16. J.G.Wilkes, D.A.Buzatu, D.J.Dare, Y.P.Dragan, M.P.Chiarelli, R.D.Holland, M.Beaudoin, T.M.Heinze, R.Nayak and A.A.Shvartsbrg, Improved cell typing by charge-state deconvolution of matrix-assisted laser desorption/ionization mass spectra, *Rapid Comm. Mass Spectrom.* **20**, 1595-1603 (2006).
17. T.Cain, D.M.Lubman and W.J.Jr.Weber, Differentiation of bacteria using protein profiles from matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Rapid Comm. Mass Spectrom.* **8**, 1026-1030 (1996).
18. M.A.Claydon, S.N.Davey, V.Edwards-Jones and D.B.Gordon, The rapid identification of intact microorganisms using mass spectrometry, *Nat. Biotechnol.* **14**, 1584-1586 (1996).
19. J.G.W.Holland, Rapid Identification of Intact Whole Bacteria Based on Spectral Patterns using Matrix-assisted Laser Desorption/Ionization with Time-of-flight Mass Spectrometry, *Rapid Comm. Mass Spectrom.* **10**, 1227-1232 (1996).
20. T.Krishnamurthy, P.L.Ross and U.Rajamani, Detection of pathogenic and non-pathogenic bacteria by matrix assisted laser desorption/ionization time of flight mass spectrometry, *Rapid Comm. Mass Spectrom.* **10**, 883-888 (1996).
21. X.Liang, K.Zheng, M.G.Qian and D.M.Lubman, Determination of bacterial protein profiles by matrix-assisted laser desorption/ionization mass spectrometry with high-performance liquid chromatography, *Rapid Comm. Mass Spectrom.* **10**, 1219-1226 (1996).
22. T.Krishnamurthy and P.L.Ross, Rapid identification of bacteria by direct matrix-assisted laser desorption/ionization mass spectrometric analysis of whole cells, *Rapid Comm. Mass Spectrom.* **10**, 1992-1996 (1996).
23. D.J.Evason, M.A.Claydon and D.B.Gordon, Exploring the limits of bacterial identification by intact cell-mass spectrometry, *J. Am. Soc. Mass Spectrom.* **12**, 49-54 (2001).
24. M.Erhard, H.von Döhren and P.Jungblut, Rapid typing and elucidation of new secondary metabolites of intact cyanobacteria using MALDI-TOF mass spectrometry, *Nat. Biotechnol.* **15**, 906-909 (1997).
25. M.Welker and M.Erhard, Consistency between chemotyping of single filaments of Planktothrix rubescens (cyanobacteria) by MALDI-TOF and the peptide patterns of strains determined by HPLC-MS, *J. Mass Spectrom.* **42**, 1062-1068 (2007).
26. C.Cirulli, G.Marino and A.Amoresano, Membrane proteome in *Escherichia coli* probed by MS3 mass spectrometry: a preliminary report, *Rapid Comm. Mass Spectrom.* **21**, 2389-2397 (2007).
27. V.Ryzhov and C.Fenselau, Characterization of the protein subset desorbed by MALDI from whole bacterial cells, *Anal. Chem.* **73**, 746-750 (2001).
28. R.J.Arnold and J.P.Reilly, Observation of *Escherichia coli* ribosomal proteins and their posttranslational modifications by mass spectrometry, *Anal. Biochem.* **269**, 105-112 (1999).
29. P.A.Demirev, J.S.Lin, F.J.Pineda and C.Fenselau, Bioinformatics and mass spectrometry for microorganism identification: proteome-wide post-translational modofications and database search algorithms for characterization of intact *H. pylori*, *Anal. Chem.* **73**, 4566-4573 (2001).
30. B.Arniri-Eliasi and C.Fenselau, Characterization of protein biomarkers desorbed by MALDI from whole fungal cells, *Anal. Chem.* **73**, 5228-5231 (2001).
31. C.K.Fagerquist, W.G.Miller, L.A.Harden, A.H.Bates, W.H.Vensel, G.Wang and R.E.Mandrell, Genomic and proteomic identification of a DNA-binding protein used in the "fingerprinting" of *Campylobacter* species and strains by MALDI-TOF-MS protein biomarker analysis, *Anal. Chem.* **77**, 4897-4907 (2005).
32. L.Sun, K.Teramoto, H.Sato, M.Torimura, H.Tao and T.Shintani, Characterization of ribosomal proteins as biomarkers for matrix-assisted laser desorption/ionization mass spectral identification of *Lactobacillus plantarum*, *Rapid Comm. Mass Spectrom.* **20**, 3789-3798 (2006).
33. K.Teramoto, H.Sato, L.Sun, M.Torimura, H.Tao, H.Yoshikawa, Y.Hotta, A.Hosoda and H.Tamura, Phylogenetic classification of *Pseudomonas putida* strains by MALDI-MS using ribosomal subunit proteins as biomarkers, *Anal. Chem.* **79**, 8712-8719 (2007).
34. J.A.Karty, S.Lato and J.P.Reilly, Detection of the bacteriological sex factor in *E. coli* by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Rapid Comm. Mass Spectrom.* **12**, 625-629 (1998).
35. Z.Wang, L.Russon, L.Li, D.C.Roser and S.R.Long, Investigation of spectral reproducibility in direct analysis of bacteria proteins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Rapid Comm. Mass Spectrom.* **12**, 456-464 (1998).
36. T.L.Williams, D.Andrzejewski, J.O.Lay and S.M.Musser, Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells, *J. Am. Soc. Mass Spectrom.* **14**, 342-351 (2003).
37. J.Ramirez and C.Fenselau, Factors contributing to peak broadening and mass accuracy in the characterization of intact spores using matrix-assisted laser desorption/ionization coupled with time-of-flight mass spectrometry, *J. Mass Spectrom.* **36**, 929-936 (2001).
38. S.C.Wunschel, K.H.Jarman, C.E.Petersen, N.B.Valentine, K.L.Wahl, D.Schauki, J.Jackman, C.P.Nelson and E.White, Bacterial analysis by

MALDI-TOF mass spectrometry: an inter-laboratory comparison, *J. Am. Soc. Mass Spectrom.* **16**, 456-462 (2005).

39. R.J.Arnold, J.A.Karty, A.D.Ellington and J.P.Reilly, Monitoring the growth of a bacteria culture by MALDI-MS of whole cells, *Anal. Chem.* **71**, 1990-1996 (1999).
40. N.B.Valentine, S.C.Wunschel, D.S.Wunschel, C.E.Petersen and K.L.Wahl, Effect of culture conditions on microorganism identification by matrix-assisted laser desorption ionization mass spectrometry, *Appl. Environ. Microbiol.* **71**, 58-64 (2005).
41. D.S.Wunschel, E.A.Hill, J.S.McLean, K.H.Jarman, Y.A.Gorby, N.B.Valentine and K.L.Wahl, Effects of varied pH, growth rate and temperature using controlled fermentation and batch culture on Matrix Assisted Laser Desorption/Ionization whole cell protein fingerprints, *J. Microb. Meth.* **62**, 259-271 (2005).
42. Y.Hathout, B.Setlow, R.M.Cabrera-Martinez, C.Fenselau and P.Setlow, Small, acid-soluble proteins as biomarkers in mass spectrometry analysis of *Bacillus* spores, *Appl. Environ. Microbiol.* **69**, 1100-1107 (2003).
43. P.A.Demirev, J.Ramirez and C.Fenselau, Tandem mass spectrometry of intact proteins for characterization of biomarkers from *Bacillus cereus* T spores, *Anal. Chem.* **73**, 5725-5731 (2001).
44. E.Elhanany, R.Barak, M.Fisher, D.Kobiler and Z.Altboum, Detection of specific *Bacillus anthracis* spore biomarkers by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Rapid Comm. Mass Spectrom.* **15**, 2110-2116 (2001).
45. J.R.Whiteaker, B.Warscheid, P.Pribil, Y.Hathout and C.Fenselau, Complete sequences of small acid-soluble proteins from Bacillus globigii, *J. Mass Spectrom.* **39**, 1113-1121 (2004).
46. P.A.Pribil, E.Patton, G.Black, V.Doroshenko and C.Fenselau, Rapid characterization of *Bacillus* spores targeting species-unique peptides produced with an atmospheric pressure matrix-assisted laser desorption/ionization source, *J. Mass Spectrom.* **40**, 464-474 (2005).
47. E.R.Castanha, A.Fox and K.F.Fox, Rapid discrimination of Bacillus anthracis from other members of the B. cereus group by mass and sequence of "intact" small acid soluble proteins (SASPs) using mass spectrometry, *J. Microbiol. Meth.* **67**, 230-240 (2006).
48. E.C.Lynn, M.-C.Chung, W.-C.Tsai and C.-C.Han, Identification of Enterobacteriaceae bacteria by direct matrix-assisted laser desorptiom/ionization mass spectrometric analysis of whole cells, *Rapid Comm. Mass Spectrom.* **13**, 2022-2027 (1999).
49. T.J.Siegrist, P.D.Anderson, W.H.Huen, G.T.Kleinheinz, C.M.McDermott and T.R.Sandrin, Discrimination and characterization of environmental strains of Escherichia coli by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS), *J. Microb. Meth.* **68**, 554-562 (2007).
50. S.Rupf, K.Breitung, W.Schellenberger, K.Merte, S.Kneist and K.Eschrich, Differentiation of mutans streptococci by intact cell matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Oral Microbiology and Immunology* **20**, 267-273 (2005).
51. H.Moura, A.R.Woolfitt, M.G.Carvalho, A.Pavlopoulos, L.M.Teixeira, G.A.Satten and J.R.Barr, MALDI-TOF mass spectrometry as a tool for differentiation of invasive and noninvasive Streptococcus pyogenes isolates, *FEMS Immun. Med. Microbiol.* **in press**, (2008).
52. C.Friedrichs, A.C.Rodloff, G.S.Chhatwal, W.Schellenberger and K.Eschrich, Rapid identification of viridans streptococci by mass spectrometric discrimination, *J. Clin. Microbiol.* **45**, 2392-2397 (2007).
53. E.Carbonnelle, J.L.Beretti, S.Cottyn, G.Quesne, P.Berche, X.Nassif and A.Ferroni, Rapid identification of staphylococci isolated in clinical microbiology laboratories by matrix-assisted laser desorption ionization - Time of flight mass spectrometry, *J. Clin. Microbiol.* **45**, 2156-2161 (2007).
54. L.Rajakaruna, G.Hallas, L.Molenaar, D.Dare, H.Sutton, V.Encheva, R.Culak, I.Innes, G.Ball, A.M.Sefton, M.Eydmann, A.M.Kearns and H.N.Shah, High trhoughput identification of clinical isolates of Staphylococcus aureus using MALDI-TOF MS of intact cells, *Infect. Gen. Evol.* **in press**, (2009).
55. V.Edwards-Jones, M.A.Claydon, D.J.Evason, J.Walker, A.J.Fox and D.B.Gordon, Rapid discrimination between methicillin-sensitive and methicillin-resistant *Staphylococcus aureus* by intact cell mass spectrometry, *J. Med. Microbiol.* **49**, 295-300 (2000).
56. Z.Du, R.Yang, Z.Guo, Y.Song and J.Wang, Identification of *Staphylococcus aureus* and determination of its methicillin resistance by matrix-assisted laser desorption/ionization time-of-flight mass Spectrometry, *Anal. Chem.* **74**, 5487-5491 (2002).
57. C.J.Keys, D.J.Dare, H.Sutton, G.Wells, M.Lunt, T.McKenna, M.McDowall and H.N.Shah, Compilation of a MALDI-TOF mass spectral database for the rapid screening and characterisation of bacteria implicated in human infectious diseases, *Infection Genetics and Evolution* **4**, 221-242 (2004).
58. M.C.Enright, D.A.Robinson, G.Randle, E.J.Feil, H.Grundmann and B.G.Spratt, The evolutionary history of methicillin-resistant Staphylococcus aureus (MRSA), *Proc. Natl. Acad. Sci. U. S. A* **99**, 7687-7692 (2002).
59. E.Seibold, R.Bogumil, S.Vorderwulbecke, S.Al Dahouk, A.Buckendahl, H.Tomaso and W.Splettstoesser, Optimized application of surface-enhanced laser desorption/ionization time-of-flight MS to differentiate *Francisella tularensis* at the level of subspecies and individual strains, *FEMS Immunol. Med. Microbiol.* **49**, 364-373 (2007).
60. J.M.Hettick, M.L.Kashon, J.E.Slaven, Y.Ma, J.P.Simpson, P.D.Siegel, G.N.Mazurek and D.N.Weissman, Discrimination of intact mycobacteria at the strain level: A combined MALDI-TOF MS and biostatistical analysis, *Proteomics* **6**, 6416-6425 (2006).
61. P.Lasch, H.Nattermann, M.Erhard, M.Stämmler, R.Grunow, N.Bannert, B.Appel and D.Naumann, A MALDI-ToF mass spectrometry compatible inactivation method for highly pathogenic microbial cells and spores, *Anal. Chem.* **80**, 2026-2034 (2008).
62. O.Schmidt and W.Kallow, Differentiation of indoor wood decay fungi with MALDI-TOF mass spectrometry, *Holzforschung* **59**, 374-377 (2005).
63. M.Erhard, U.C.Hipler, A.Burmester, A.A.Brakhage and J.Wöstemeyer, Identification of dermatophyte species causing onychomycosis and tinea pedis by

MALDI-TOF mass spectrometry, *Exp. Dermatol.* **17**, 365-371 (2008).

64. F.Seyfarth, M.Ziemer, M.Kaatz, H.Sayer, A.Burmester, M.Erhard, M.Welker, S.Schliemann, E.Strabe and U.C.Hipler, Case report: The use of ITS DNA sequence analysis and MALDI-TOF mass spectrometry in diagnosing an infection with *Fusarium proliferatum*, *Exp. Dermatol.* **17**, 965-971 (2008).
65. E.N.Villegas, S.T.Glassmeyer, M.W.Ware, S.L.Hayes and F.W.Schaefer, Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry-based analysis of Giardia lamblia and Giardia muris, *Journal of Eukaryotic Microbiology* **53**, S179-S181 (2006).
66. M.L.Magnuson, J.H.Owens and C.A.Kelty, Characterization of *Cryptosporidium parvum* by matrix-assisted laser desorptio/ionisation time-of-flight mass spectrometry, *Appl. Environ. Microbiol.* **66**, 4720-4724 (2000).
67. E.Vanlaere, K.Sergeant, P.Dawyndt, W.Kallow, M.Erhard, H.Sutton, D.Dare, B.Samyn, B.Devreese and P.Vandamme, Matrix assisted laser desorption ionization-time of flight mass spectrometry of intact cells allows rapid identification of *Burkholderia cepacia* complex species, *J. Microb. Meth.* **75**, 279-286 (2008).
68. R.J.Arnold and J.P.Reilly, Fingerprint matching of *E. coli* strains with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of whole cells using a modified correlation approach, *Rapid Comm. Mass Spectrom.* **12**, 630-636 (1998).
69. P.A.Demirev, Y.P.Ho, V.Ryzhov and C.Fenselau, Microorganism identification by mass spetrometry and protein database searches, *Anal. Chem.* **71**, 2732-2738 (1999).
70. K.H.Jarman, D.S.Daly, C.E.Peterson, A.J.Saenz, N.B.Valentine and K.L.Wahl, Extracting and visualizing matrix-assisted laser desorption/ionization time-of-flight mass spectral fingerprints, *Rapid Comm. Mass Spectrom.* **13**, 1586-1594 (1999).
71. K.H.Jarman, S.T.Cebula, A.J.Saenz, C.E.Petersen, N.B.Valentine, M.T.Kingsley and K.L.Wahl, An Algorithm for Automated Bacterial Identification Using Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry, *Anal. Chem.* **72**, 1217-1223 (2000).
72. J.J.Bright, M.A.Claydon, M.Soufian and D.B.Gordon, Rapid typing of bacteria using matrix-assisted laser desorption ionisation time-of-flight mass spectrometry and pattern recognition software, *J. Microbiol. Methods* **48**, 127-138 (2002).
73. S.-Y.Hsieh, C.-L.Tseng, Y.-S.Lee, A.-J.Kuo, C.-F.Sun, Y.-H.Lin and J.-K.Chen, Highly efficient classification and identification of human pathogenic bacteria by MALDI-TOF MS, *Molecular Cellular Proteomics* **7**, 448-456 (2008).
74. M.A.Domin, K.J.Welham and D.S.Ashton, The effect of solvent and matrix combinations on the analysis of bacteria by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry, *Rapid Comm. Mass Spectrom.* **13**, 222-226 (1999).
75. D.J.Evason, M.A.Claydon and D.B.Gordon, Effects of ion mode and matrix additives in the identification of bacteria by intact cell mass spectrometry, *Rapid Comm. Mass Spectrom.* **14**, 669-672 (2000).
76. K.H.Jarman, D.S.Daly, K.K.Anderson and K.L.Wahl, A new approach to automated peak detection, *Chemometrics and Intelligent Laboratory Systems* **69**, 61-76 (2003).
77. Z.Wang, K.Dunlop, S.R.Long and L.Li, Mass spectrometric methods for generation of protein mass database used for bacterial identification, *Anal. Chem.* **74**, 3174-3182 (2002).
78. S.Sauer, A.Freiwald, T.Maier, M.Kube, R.Reinhardt, M.Kostrzewa and K.Geider, Classifiction and identification of bacteria by mass spectrometry and computional analysis, *PLoS One* **3**, e2843 (2008).
79. W.Kallow, R.Dieckmann, N.Kleinkauf, M.Erhard and T.Neuhof, *Method of identifying microorganisms using MALDI-TOF-MS*, European Patent, (2000).
80. W.Pusch, Bruker Daltronics: leading the way from basic research to mass-spectrometry-based clinical applications, *Pahrmacogenomics* **8**, 663-668 (2007).
81. J.Gielen, M.Erhard, W.Kallow, M.Krönke and O.Krut. Rapid pathogen identification by MALDI-TOF mass spectrometry/SARAMIS database in clinical microbiological routine diagnostics. ECCMID 17 . 2007. Munich, Germany.
82. E.Stackebrandt, W.Frederiksen, G.M.Garrity, P.A.D.Grimont, P.Kämpfer, M.C.J.Maiden, X.Nesme, R.Rosselló-Mora, J.Swings, H.G.Trüper, L.Vauterin, A.C.Ward and W.B.Whiotman, Report of the *ad hoc* committee for the re-evaluation of the species definition in bacteriology, *Int. J. Syst. Evol. Microbiol.* **52**, 1043-1047 (2002).
83. G.M.Pupo, R.Lan and P.R.Reeves, Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics, *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10567-10572 (2000).
84. R.Lan, M.C.Alles, K.Donohhoe, M.B.Martinez and P.R.Reeves, Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* ssp., *Infect. Immun.* **72**, 5080-5088 (2004).
85. A.Mellmann, J.Cloud, T.Maier, U.Keckevoet, I.Rarnminger, P.Iwen, J.Dunn, G.Hall, D.Wilson, P.R.Lasala, M.Kostrzewa and D.Harmsen, Evaluation of Matrix-Assisted Laser Desorption Ionization-Time-of-Flight mass spectrometry in comparison to 16S rRNA gene sequencing for species identification of nonfermenting bacteria, *Journal of Clinical Microbiology* **46**, 1946-1954 (2008).
86. P.Krader and D.Emerson, Identification of archaea and some extremophilic bacteria using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry, *Extremophiles* **8**, 259-268 (2004).
87. M.J.Donohue, A.W.Smallwood, S.Pfaller, M.Rodgers and J.A.Shoemaker, The development of a matrix-assisted laser desorption/ionization mass spectrometry-based method for the protein fingerprinting and identification of Aeromonas species using whole cells, *J. Microbiol. Meth.* **65**, 380-389 (2006).
88. P.A.Pribil and C.Fenselau, Characterization of Enterobacteria using MALDI-TOF mass spectrometry, *Anal. Chem.* **77**, 6092-6095 (2005).
89. J.E.Camara and F.A.Hays, Discrimination between wild-type and ampicillin-resistant Escherichia coli by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Analytical and Bioanalytical Chemistry* **389**, 1633-1638 (2007).
90. C.Y.Pierce, J.R.Barr, A.R.Woolfitt, H.Moura, E.I.Shaw, H.A.Thompson, R.F.Massung and F.M.Fernandez, Strain and phase identification of

the US category B agent *Coxiella burnetii* by matrix assisted laser desorption/ionization time-of-flight mass spectrometry and multivariate pattern recognition, *Analytica Chimica Acta* **583**, 23-31 (2007).

91. A.Schaller, R.Troller, D.Molina, S.Gallati, C.Aebi and P.S.Meier, Rapid typing of *Moraxelia catarrhalis* subpopulations based on outer membrane proteins using mass spectrometry, *Proteomics* **6**, 172-180 (2006).
92. A.M.Haag, S.N.Taylor, K.H.Johnston and R.B.Cole, Rapid identification and speciation of *Haemophilus* bacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *J. Mass Spectrom.* **33**, 750-756 (1998).
93. C.K.Fagerquist, E.Yee and W.G.Miller, Composite sequence proteomic analysis of protein biomarkers of Campylobacter coli, C-lari and C-concisus for bacterial identification, *Analyst* **132**, 1010-1023 (2007).
94. R.E.Mandrell, L.A.Harden, A.Bates, W.G.Miller, W.F.Haddon and C.K.Fagerquist, Speciation of *Campylobacter coli*, *C. jejuni*, *C. helveticus*, *C. lari*, *C. sputorum*, and *C. upsaliensis* by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Appl. Environ. Microbiol.* **71**, 6292-9307 (2005).
95. C.K.Fagerquist, A.H.Bates, S.Heath, B.C.King, B.R.Garbus, L.A.Harden and W.G.Miller, Subspeciating Campylobacter jejuni by proteomic analysis of its protein biomarkers and their post-translational modifications, *Journal of Proteome Research* **5**, 2527-2538 (2006).
96. C.L.Nilsson, Fingerprinting of *Helicobacter pylori* strains by matrix-assisted laser desorption/ionization mass spectrometric analysis, *Rapid Comm. Mass Spectrom.* **13**, 1067-1071 (1999).
97. M.Vargha, Z.Takats, A.Konopka and C.H.Nakatsu, Optimization of MALDI-TOF MS for strain level differentiation of *Arthrobacter* isolates, *J Microbiol Methods* **66**, 399-409 (2006).
98. M.Pignone, K.M.Greth, J.Cooper, D.Emerson and J.Tang, Identification of mycobacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *J. Clin. Microbiol.* **44**, 1963-1970 (2006).
99. M.J.Stump, G.Black, A.Fox, K.F.Fox, C.E.Turick and M.Matthews, Identification of marker proteins for Bacillus anthracis using MALDI-TOF IVIS and ion trap MS/MS after direct extraction or electrophoretic separation, *J. Separ. Sci.* **28**, 1642-1647 (2005).
100. Y.Hathout, P.A.Demirev, Y.P.Ho, J.L.Bundy, V.Rhyzov, L.Sapp, J.Stutler, J.Jackman and C.Fenselau, Identification of *Bacillus* spores by matrix-assisted laser desorption ionization-mass spectrometry, *Appl. Environ. Microbiol.* **65**, 4313-4319 (1999).
101. V.Ryzhov, Y.Hathout and C.Fenselau, Rapid characterization of spores of Bacillus cereus group bacteria by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry, *Appl. Environ. Microbiol.* **66**, 3828-3834 (2000).
102. S.Barbuddhe, T.Maier, G.Schwarz, M.Kostrzewa, E.Domann, T.Chakraborty and T.Hain, Rapid identification of listeria species using MALDI-TOF MS fingerprinting with MALDI BioTyper, *International Journal of Medical Microbiology* **297**, 18-19 (2007).
103. P.A.Majcherczyk, T.McKenna, P.Moreillon and P.Vaudaux, The discriminatory power of MALDI-TOF mass spectrometry to differentiate between isogenic teicoplanin-susceptible and teicoplanin-resistant strains of methicillin-resistant Staphylococcus aureus, *FEMS Microbiol. Lett.* **255**, 233-239 (2006).
104. R.A.Giebel, W.Fredenberg and T.R.Sandrin, Characterization of environmental isolates of Enterococcus spp. by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Water Research* **42**, 931-940 (2008).
105. K.J.Welham, M.A.Domin, K.Johnson, L.Jones and D.S.Ashton, Characterization of fungal spores by laser desorption/ionization time-of-flight mass spectrometry, *Rapid Comm. Mass Spectrom.* **14**, 307-310 (2000).
106. P.A.Demirev, A.B.Feldman, D.Kongkasuriyachai, P.Scholl, D.Sullivan and N.Kumar, Detection of malaria parasites in blood by laser desorption mass spectrometry, *Anal. Chem.* **74**, 3262-3266 (2002).
107. F.W.F.Lee, K.C.Ho and S.C.L.Lo, Rapid identification of dinoflagellates using protein profiling with matrix-assisted laser desorption/ionization mass spectrometry, *Harmful Algae* **7**, 551-559 (2008).
108. M.R.Perera, V.A.Vanstone and M.G.K.Jones, A novel approach to identify plant parasitic nematodes using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Rapid Comm. Mass Spectrom.* **19**, 1454-1460 (2005).
109. M.R.Perera, R.D.F.Vargas and M.G.K.Jones, Identification of aphid species using protein profiling and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Entomologia Experimentalis et Applicata* **117**, 243-247 (2005).

Table 1: Compilation of studies on mass spectral analysis of microbial taxa and their potential to differentiate closely related species based on mass spectral fingerprints.

Taxon	Comment	ref.
Prokaryotes		
Archeae	characterization of archeae and extremophil bacteria	86
Eubacteria		
Aeromonas sp.	genus-specific mass signals and discrimination of species	87
Enterobacteriaceae	family- and genus-specific biomarkers	48
Enterobacteriaceae	analysis after on-slide treatment with trypsin	88
E. coli	discrimination of ampicillin resistant and susceptible strains	89
E. coli	differentiation of environmental isolates by host origin	49
Erwinia sp.	typing of subspecies	78
Yersinia sp.	mass spectral typing after inactivation with 80% TFA	61
Coxiella burnetii	identification in different growth phases	90
Moraxella catarrhalis	mass spectral typing consistent with genotyping	91
Francisella tularensis	SELDI-TOF MS discrimination of subspecies	59
Pseudomonas putida	mass spectral classification consistent with genotyping	33
Haemophilus sp.	discrimination of pathogenic and non-pathogenic species	92
Burkholderia sp.	discrimination of *B. cepacia*-complex species	67
Campylobacter sp.	characterization of protein biomarkers in three species	93
Campylobacter sp.	discrimination of six species; identification of biomarkers	94
Campylobacter jejuni	subspeciation; identification of biomarkers	95
Helicobacter pylori	species- and strain-specific biomarkers	96
Arthrobacter sp.	strain level differentiation	97
Mycobacterium sp.	characterization and identification of species and strains	60
Mycobacterium sp.	discrimination of 13 species	98
Bacillus anthracis	discrimination from other species of *B. cereus* group	47
Bacillus anthracis	biomarkers in spores	44
Bacillus anthracis	characterisation of biomarkers by IT MS/MS	99
Bacillus sp.	biomarkers in spores; strain-specific in *B. cereus*	100
Bacillus sp.	identification of SASPs in spores as biomarkers	42
Bacillus sp.	mass spectral typing after inactivation of spores with TFA	61
Bacillus cereus	pretreatment with corona plasma discharge	101
Listeria sp.	differentiation of clonal lineages	102
Staphylococcus	discrimination of 23 species and subspecies	53
Staphylococcus	discrimination of MRSA and MSSA	56
Staphylococcus	discrimination of MRSA and MSSA	55
Staphylococcus	high throughput identification of hospital isolates	54
Staphylococcus	MRSA differentiation with respect to teicoplanin resistance	103
Streptococcus viridans gr.	identification of clinical isolates (excluding *Str. Pneumoniae*)	52
Streptococcus mutans gr.	determination of species of clinical isolates	50
Streptococcus pyogenes	differentiation of invasive and non-invasive isolates	51
Enterococcus sp.	source tracking of environmental isolates	104
Lactobacillus plantarum	identification of ribosomal proteins as biomarkers	32
Nonfermenter	identification of clinical isolates	85
Eukaryota		
Fungi		
Fusarium sp.	discrimination of *Fusarium* species	64
Serpula sp.	identification of indoor wood-decay fungi	62
dermatophytes	identifcation of clinical *Trichophyton* isolates	63
fungal spores	reproducibility of mass fingerprints	105
Protists		
Plasmodium sp.	detection in processed blood samples	106
cryptosporidia	specific biomarkers for two species	66
Giardia lamblia	discrimination of species	65
dinoflagellates	identification of HAB species	107
Metazoa		
nematoda	identification of plant parasites	108
aphids insecta	discrimination of species	109

Table 2: Summary of the evaluation of MALDI-TOF MS/SARAMIS identification of isolates from urine samples compared to identification obtained using VITEK II. (Studies carried out in the laboratory of Dr. Stein, Mönchengladbach, Germany). Numbers for correct identifications with different confidence levels (conf.) are given separately. Numbers in parentheses refer to percentages. For details see text.

	n	>99-90 % conf.	90-80 % conf.	no ID	false ID	no spectrum
Enterobacteriaceae	338	319 (94.4)	12 (3.6)	6 (1.8)	0 (0)	1 (0.3)
Enterococci & Streptococci	84	70 (83.3)	8 (9.5)	3 (3.6)	0 (0)	3 (3.6)
Staphylococci	79	61 (77.2)	12 (15.2)	4 (5.1)	0 (0)	2 (2.5)
Nonfermenter	68	64 (94.1)	4 (5.9)	0 (0)	0 (0)	0 (0)
Total	569	514 (90.3)	36 (6.3)	13 (2.3)	0 (0)	6 (1.1)

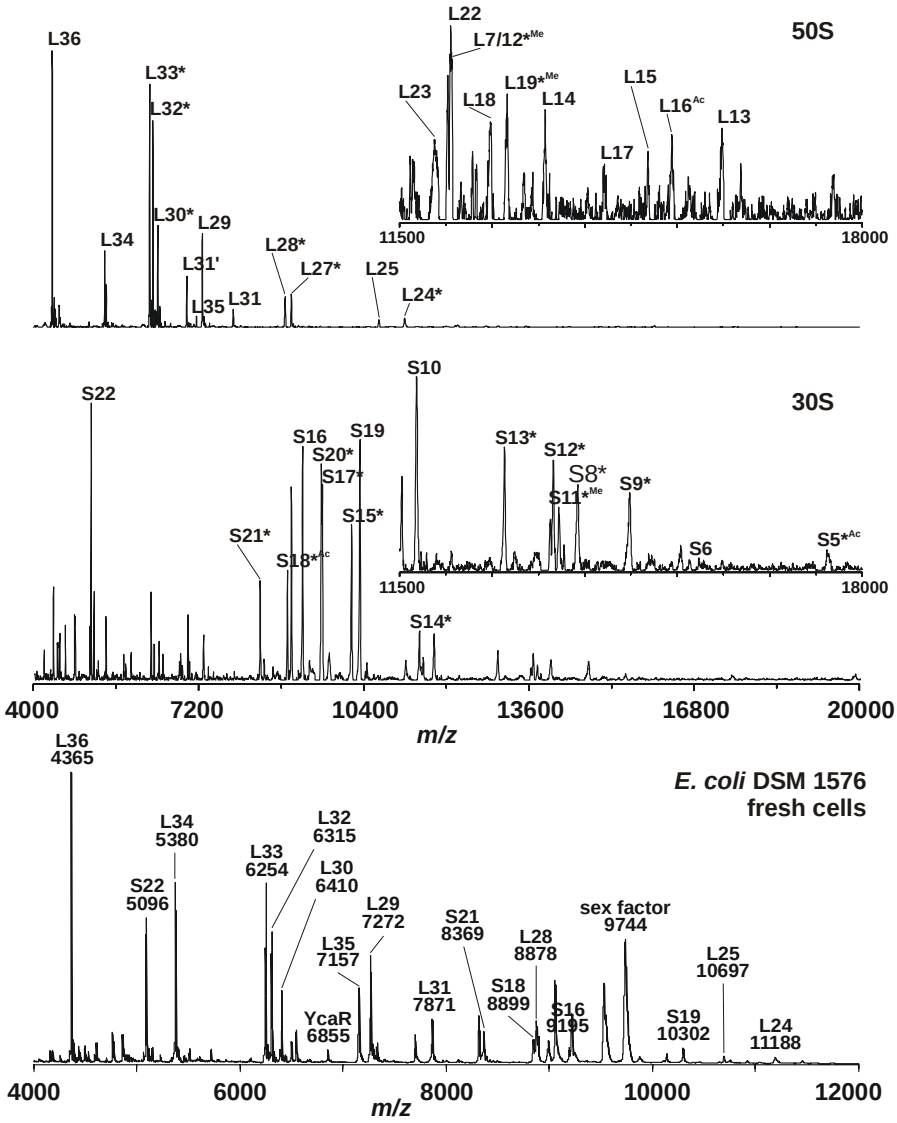

Figure 1: MALDI-TOF mass spectra of isolated ribosomes of *Escherichia coli* K12 and of fresh cells directly prepared on the sample target. Mass spectra of 50S (top) and 30S (middle) ribosomal subunits were recorded on a Voyager DE Pro system in 2000 while fresh cells (bottom) were analysed on a Shimadzu AXIMA CFRplus in 2007. Ribosomal proteins are indicated by L (large 50S subunit) or S (small 30S subunit) and numbers. In the two upper plots, the inserts show an enlarged mass range as indicated. Asterisks indicate an N-terminal methionine cleavage; superscipt 'Me' a single methylation, and superscript 'Ac' an acetylation [28]. YcaR refers to a non-characterized protein and 'sex factor' to an F-plasmid related protein [34].

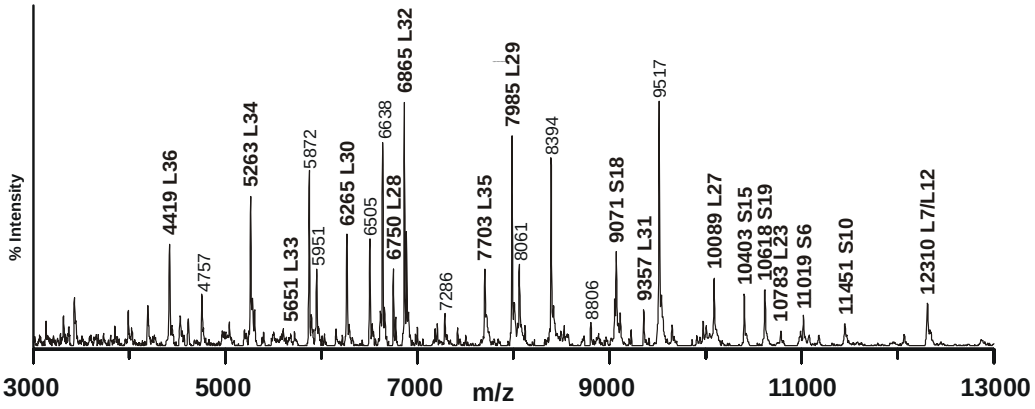

Figure 2: Whole cell MALDI-TOF mass spectrum of a clinical isolate of *Streptococcus pneumoniae*. Ribosomal proteins are indicated by their numbers referring to the small 30S (S) or the large 50S subunit (L) proteins. Assignment was done based on comparison of protein masses obtained from translated genomic sequences of multiple *Str. pneumoniae* strains available from SwissProt website and taken into consideration possible posttranslational modifications by N-terminal methionine cleavage and/or mono-methylation.

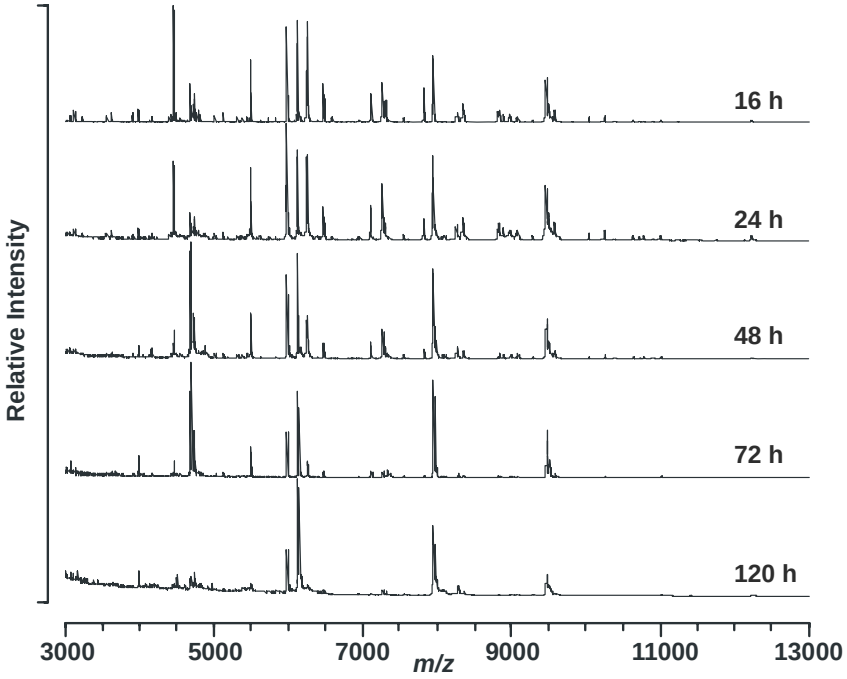

Figure 3: Whole cell MALDI-TOF mass spectra of *Proteus mirabilis* DSM 4479 acquired after cultivation times as indicated. The strain was grown on Columbia blood agar at 37° C and shows a remarkably stable profile in spite of gross changes taking place within the cell during various phases of growth.

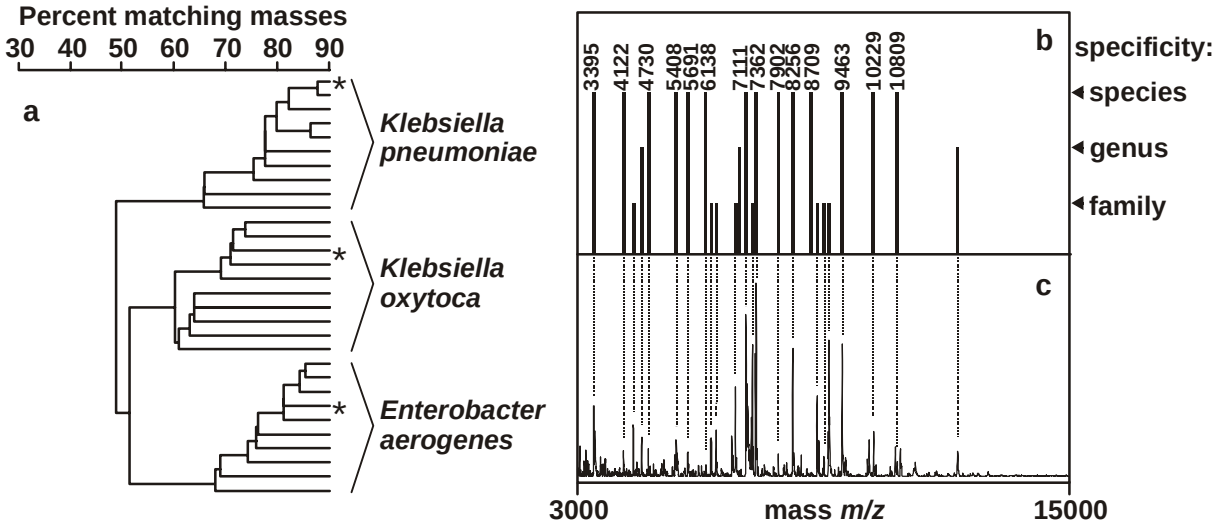

Figure 4: Dendrogram of mass spectral fingerprints of ten strains each of three closely related species within the diverse family Enterobacteriaceae (a), illustration of a SARAMIS SuperSpectrum for *Klebsiella oxytoca* with taxonomic specificities of mass signals indicated by their height (b), and a mass spectrum of a clinical isolate of *K. oxytoca* with mass signals matching to the SuperSpectrum indicated by dotted lines (c). Asterisks indicate DSM reference strains of respective species. For the calculation of SuperSpectra see text.