



**HAL**  
open science

## Pathogen proteotyping: A rapidly developing application of mass spectrometry to address clinical concerns

Lucia Grenga, Olivier Pible, J. Armengaud

### ► To cite this version:

Lucia Grenga, Olivier Pible, J. Armengaud. Pathogen proteotyping: A rapidly developing application of mass spectrometry to address clinical concerns. *Clinical Mass Spectrometry*, 2019, 14, pp.9-17. 10.1016/j.clinms.2019.04.004 . hal-02886103

**HAL Id: hal-02886103**

**<https://hal.inrae.fr/hal-02886103>**

Submitted on 20 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Pathogen proteotyping: a rapidly developing application of mass spectrometry to address clinical concerns

Lucia Grenga<sup>a\*</sup>, Olivier Pible<sup>a</sup> and Jean Armengaud<sup>a</sup>

<sup>a</sup> Laboratoire Innovations technologiques pour la Détection et le Diagnostic (Li2D), Service de Pharmacologie et Immunoanalyse (SPI), CEA, INRA, F-30207 Bagnols sur Cèze, France.

\* Corresponding author: Lucia Grenga, CEA-Marcoule, DRF-Li2D, Laboratory “Innovative technologies for Detection and Diagnostics”, BP 17171, F-30200 Bagnols-sur-Cèze, France; [lucia.grenga@cea.fr](mailto:lucia.grenga@cea.fr); Tel: +00 33 4 66 79 54 01; Fax: +00 33 4 66 79 19 05.

Running title: Pathogen proteotyping

## **Abstract**

For the rapid and reliable differentiation of clinically-relevant bacterial species, mass spectrometry-based methods have emerged in recent years as valid alternatives to existing techniques. Mass profiles generated by whole-cell Matrix-Assisted Laser Desorption Ionization-Time of Flight mass spectrometry have revolutionized microorganism identification and proven their potential for proteotyping at the species level. Indeed, the methodology has been widely deployed in clinical settings. However, the low resolution and dynamic range of the methodology has limited its capacity to distinguish between subspecies. This discrimination capacity is pivotal in cases where certain strains display virulence or antibiotic resistance, and for epidemiologic analyses. Moreover, sensitivity and specificity are both key parameters when attempting to discriminate between microorganisms present in complex multi-pathogenic samples. These two parameters are also essential to meet the growing interest in the characterization of microorganisms contained within even more complex samples, such as the human microbiome. Tandem mass spectrometry, with its high resolution, holds great potential for use in the real-time direct analysis of pathogens at the most relevant taxonomic rank in routine clinical practice. This review explores the numerous benefits and challenges of implementing advanced proteotyping methods, based on tandem mass spectrometry, in clinical laboratories. We provide an overview of the current applications and methodologies, while also discussing recent improvements and potential new approaches for typing, as well as their future applications.

**Keywords:** Proteotyping, Mass spectrometry, MALDI-TOF, Tandem mass spectrometry, Pathogen typing

## **Abbreviations:**

FTICR - Fourier transform ion cyclotron resonance

MALDI-TOF - Matrix-Assisted Laser Desorption Ionization-Time Of Flight

MLST - multilocus sequence typing  
MS/MS - tandem mass spectrometry  
PFGE - pulsed-field gel electrophoresis

## **1. Introduction**

### *1.1 Proteotyping: clarifying the concept.*

Literally, 'typing by proteins', proteotyping refers to the use of proteins as distinguishing factors for taxonomical purposes. This is a relatively recent concept as, traditionally, taxonomists rely on molecular typing methods mostly based on DNA restriction or amplification, e.g. pulsed-field gel electrophoresis (PFGE) or multilocus sequence typing (MLST) [1]. Although widely recognized, these techniques remain laborious and time-consuming to perform. As mass spectrometry (MS)-based proteomics has progressed, conventional biochemical/phenotypic methods have been replaced by more cost-effective discriminatory MS-based typing approaches, in particular by whole-cell Matrix-Assisted Laser Desorption Ionization-Time Of Flight (MALDI-TOF) MS proteotyping. Depending on the instrumentation and platforms available, proteotyping can now be achieved by analyzing mass patterns from intact proteins (top-down approaches, such as for whole-cell MALDI-TOF) or peptides generated from proteins (bottom-up approaches). Numerous innovative approaches have been developed especially to discriminate between microorganisms at species/subspecies level.

### *1.2 Functional proteotyping: gaining insights into pathogen functioning.*

Extending the *sensu stricto* definition of proteotyping, Karlsson et al. [2] recently suggested that insights into the physiological state and functions of microbial populations could be obtained by proteotyping microorganisms grown under differing culture conditions (e.g., nutrients, stress). Because of the direct link between phenotypic traits and proteome characteristics, this concept reconnects molecular information with historical phenotypic taxonomic classification. However, functional proteotyping must be distinguished from classical proteomic characterization of a single microorganism grown under different conditions to understand its physiological response through comparative analysis. Rather, proteotyping should be considered to exclusively describe the use of protein data to discriminate between groups of organisms, either in terms of taxonomy or functionality. In real terms, these approaches could be applied to type several antibiotic-resistance mechanisms or detect the presence of specific toxins and other virulence factors. However, current MS-based typing methods, such as whole-cell MALDI-TOF, rely on the mass patterns produced by small, abundant and basic proteins, without actually identifying proteins, and thus cannot be used to assess functional traits.

### *1.3 Proteotyping and metaproteomics: are the views distinct?*

Metaproteomics, a term coined more than a decade ago [3], aims to resolve and quantify the major active metabolic pathways in complex microbiota, and, thereby, establish genotype-phenotype relationships. This methodology relies on: metagenomics data, to select the most appropriate protein-sequence database; tandem MS (MS/MS), to acquire data on the proteins present; and bioinformatics, for data interpretation. Two facets can be distinguished: taxonomical characterization of the sample will allow the identification of the main taxa present, based on experimental protein data; then the identification and quantification of proteins can be used to functionally characterize the main metabolic pathways. Metaproteomics and proteotyping are clearly different, as the latter is a technology to establish which organisms are present in any given complex biological system, the list of which can then be used to answer larger questions by applying

metaproteomics techniques. For comprehensive metaproteomics, rapid and accurate characterization of highly complex samples in terms of both taxonomical purpose and functional characterization are essential, and the need for large amounts of data becomes obvious. In this context, new MS-based approaches are emerging. To glean further information with high accuracy, shotgun proteomics characterization with MS/MS measurements is essential. As shotgun proteomics provides more information, including the molecular weight of the analytes and their amino acid sequence, it outperforms whole-cell MALDI-TOF-based methods and produces meaningful and complementary information about functional bacterial traits. In this context, advances in MS technology to increase detection sensitivity and dynamic range, make it possible to resolve thousands of peptide species that vary widely in abundance within a sample, paving the way for the characterization of more complex biological systems.

This review outlines the applications of proteotyping and their usefulness in clinical practice. Moreover, current and novel methodologies are reviewed, with particular attention paid to their limitations, as well as to the recent improvements favouring applications in routine clinical practice. Finally, attractive future proteotyping applications are suggested.

## **2. Proteotyping: clinical relevance**

By offering all the advantages derived from the use of MS, such as high-speed, low-cost, simplicity, and applicability to a wide range of microbes, implementation of proteotyping in routine clinical practice has provided great advantages. It contributes to improving patient care by reducing time to diagnosis. In addition, being based on less expensive and less expertise-demanding methodologies than conventional techniques, it helps to cut the costs of clinical microbiology laboratories [4, 5]. Thanks to continuous improvements in the analytical specificity and sensitivity of mass spectrometers, numerous proteotyping MS-based methods have been proposed and, at present, some are becoming routine procedures, particularly in clinical microbiology.

Different strains from a given species may have distinct phenotypic behaviours, such as a higher capacity to cause invasive disease, to asymptotically colonize a host, or to present resistance to antimicrobials. In this scenario, proteotyping fulfils various purposes, from discrimination between epidemic clones with characteristic pathogenicity, to the comparison of clinical and environmental isolates to track the spread of specific subgroups and, for example, to control nosocomial and cross-infections [6, 7] (Fig 1). The rapidity and increased diagnostic resolution of MS-based proteotyping compared to traditional methods have become critical, in particular for severe microbial infections, such as sepsis, acute meningitis, etc. In addition, many researchers have demonstrated that MS-based approaches can be used to initiate an effective therapy early, and, thus, prevent a possible pandemic, for example by rapidly screening influenza virus subtypes [7]. MS-based strain-typing MS-based has also been described for a few *Candida* species [8, 9], although fungal proteotyping has moved at a slower pace than bacterial and viral typing. This slower pace is linked to the inherent biological complexity of fungi, and to the fact that fungal samples are less likely than bacteria to release proteins. As a result, sample preparation protocols are often labour-intensive and, up to now, only low numbers of representative taxa are found in commercially-available databases [10]. Nonetheless, proteotyping represents an expanding area for many laboratories, and there is great potential for its application and for the development of diagnostic methods [11].

A field of great significance where proteotyping could be applied is the identification of virulence and antibiotic-resistance factors for detected bacteria (Fig 1). In fact, in a world where

antibiotic-resistance rates are constantly progressing, subspecies-level identification, as well as rapid and correct determination of antibiotic-resistance patterns are increasingly needed. In this context, in addition to phenotypic diagnostic approaches [12], specific tools for approaches not reliant on prior knowledge or inferences are necessary for difficult cases, such as bacterial meningitis, respiratory tract pathogens or catheter-related bloodstream infections. Here, the increased diagnostic resolution and the shorter time-to-results offered by MS-based proteotyping, compared to traditional approaches, are hugely relevant to avoid prolonging empirical, and potentially inappropriate, antibacterial therapies [13]. Prior treatments, often relying on broad-spectrum antibiotics, can lead to unsuccessful pathogen recovery and, consequently, unavailability of drug susceptibility data. Furthermore, overtreatment with successful empirical broad-spectrum compounds is, at least partially, linked to the discrepancy between traditional microbiological procedures and the clinical need for rapid results [12]. In research laboratories, mass spectrometry has already been successfully used to detect antibiotic-resistance mechanisms through several approaches based on the analysis of antibiotics and their degradation products or rRNA modifications [14, 15]. However, a promising approach for routine testing is bacterial typing based on the direct identification of bacterial resistance determinants [13]. Through this approach, proteotyping could inform treatment decisions, facilitate assessment of the clinical relevance of microbial isolates or directly guide the selection of antimicrobials based on known patterns of intrinsic resistance and local susceptibility data. In this scenario, the potential of proteotyping, unlike the molecular tests, is not limited by the multiplicity of different resistance genes or newly emerging resistance variants [16]. Among the remarkable examples of its application, Josten et al. [17] exploited its capacity to sensitively and specifically discriminate between subgroups of nosocomial methicillin-resistant *Staphylococcus aureus* (MRSA) by detecting the small peptide PSM-mec. Similar findings were reported by Nagy et al. [18] when typing strains of the anaerobic pathogen *Bacteroides fragilis*. The detection of a specific  $m/z$  peak in the spectrum profile correlated with resistance to carbapenems, one of the most important groups of antimicrobials used to treat mixed infections. Similarly, Fagerquist et al. [19] directly distinguished *Escherichia coli* O157:H7 subtypes from Shiga toxin-producing *E. coli* by identifying Shiga toxin 2 thanks to its sequence-specific fragment ions. However, although proteotyping based on virulence or resistance determinants represents a relevant and promising tool to detect virulent isolates and antibiotic resistance, no method has yet been validated in clinical trials [4]. For application of these methods in routine clinical practice, standardized workflows or databases and software tools must be developed to provide the level of reproducibility and reliability required from a diagnostic test.

The incorporation of proteotyping into clinical routine also represents a step toward the advent of personalized medicine (Fig 1). Although still in its infancy, this methodology has considerable potential to greatly increase diagnostic accuracy, while reducing time and costs by allowing, for example, differentiation between benign colonizers and clinically-relevant findings, or to identify multiple species in co-infections. More generally, proteotyping could contribute – by providing a holistic view of the patient’s condition – to the development of personalized treatment regimens that maximize benefits and minimize adverse effects. Such tailored treatments are extremely important, for example, in immunocompromised patients for whom the treatment must be precisely selected. In this context, one of the advantages of routinely using MS-based proteotyping methods in clinical microbiology laboratories is represented by the possibility that the clinical samples can be directly analyzed without the need for culture steps. Successful identification of pathogens in the sample itself have already been demonstrated in various settings (e.g., positive blood cultures or urine samples following a concentration step) [20], suggesting that further development of protocols may allow direct typing and validation of the method for routine applications.

### 3. Current methodologies: MALDI typing methods

To date, routine clinical proteotyping relies on the analysis of protein mass spectra by means of whole-cell MALDI-TOF MS analysis. Within the last decade, MALDI-TOF MS has been fully integrated into the routine of clinical laboratories where it has revolutionized the identification of bacterial species, yeast isolates, filamentous fungi and dermatophytes [21-23]. The information acquired during routine analyses to identify isolates could be further exploited, at no additional cost, for their typing. This possibility has driven the development of novel approaches and their subsequent integration into clinical workflows.

MALDI-TOF MS is used to identify microorganisms based on mass spectral patterns obtained from whole cells following a simple procedure where the sample is mixed with a chemical matrix and deposited on a MALDI plate. A characteristic mass spectrum is acquired from the sample and then statistically compared to a database of MS patterns/fingerprints that were collected from reference samples under exactly the same conditions (Fig. 2). The identity of the microorganism is determined based on which set in the database provides the best match with the spectra obtained from the sample [11]. In addition, a quality score of the species assignment is calculated. Mass spectral patterns used for species-level identification are derived from basic abundant proteins (e.g., mainly ribosomal proteins and a few housekeeping proteins such as the HU protein), all of which are between 2 to 20 kDa. The abundance of these proteins in the cell is relatively independent of the bacterial growth state or external stimuli, but sample conditions should be identical to those used to generate the database. Slight variations in the mass of these proteins (e.g., due to amino acid substitutions deriving from non-silent mutations in their corresponding genes) in isolates within the same species could represent candidate biomarkers for subspecies discrimination [24]. For example, *Francisella tularensis subspecies tularensis*, the agent responsible for tularemia, can be differentiated from the other subspecies by two specific MALDI-TOF  $m/z$  peaks [25].

Despite several works reporting a higher discriminatory power of MALDI-TOF compared to conventional techniques, the use of MALDI typing in clinical settings is still in its infancy; until very recently, results remained controversial for several bacterial species. Further improvements have since been made to the database, mainly by including spectra from the most difficult bacterial cases, as was recently illustrated for *Brucella* [26]. Thus, although Ueda et al. [27] demonstrated that MALDI typing could be used for clone-level identification of methicillin-resistant *S. aureus* with accuracy equivalent to PFGE, Purighalla et al. [28], by highlighting differences in clustering for nosocomial *Klebsiella pneumoniae* isolates on the basis of genomic or proteomic signatures, revealed the need for further investigation to establish the widespread applicability of MALDI-TOF methods for clonality studies across a wider variety of bacteria implicated in nosocomial infections. A similar conclusion was drawn by Pinto et al. [29] following typing of *Streptococcus pneumoniae* isolates. Thus, further improvements in the performance of the method are eagerly awaited in the near future.

Although several successful examples of intraspecific group segregation have been published [23, 30], the discriminatory potential of the methodology is not the same for all pathogens [31]. Some variability in the limits of taxonomic resolution can probably be ascribed to the evidence that resolution is taxon-dependent [32] and often reflects known identification problems related to certain genera. Examples of these limits are linked to difficulties in discriminating organisms at the species level, as described for *Burkholderia mallei* / *Burkholderia pseudomallei*, and *Streptococcus mitis* / *Streptococcus oralis* / *Streptococcus pneumoniae* [33]. Below the species level, MALDI-TOF MS has been reported to have low discriminatory power for serotypes of *E. coli*, biovars of *Brucella suis* and clinically-relevant multidrug resistant *Klebsiella pneumoniae* clones [34]. To discriminate between subspecies, Gekenidis et al. [35] developed a workflow that goes beyond the classical

MALDI biotyping. Their high-intensity focused ultrasound-assisted proteomics workflow allows greater taxonomic discrimination by extending the mass range and type of potential biomarker peptides. A broader mass range (from 15 to 75 kDa) could also increase discrimination power for protein biomarkers, as shown by other groups [36-38]. Although in some cases statistical approaches could be implemented to improve the discriminative power of MALDI typing methods and identify reliable biomarkers [24], additional tools for high-resolution typing are clearly needed.

Unsuccessful identification of reliable peak biomarkers could also be explained by biological (e.g., strain sets) and methodological differences (e.g. technological issues with the mass spectrometer and/or the choice of informatics and statistical tools) [39]. For example, Fangous et al. [40] applied a classification algorithm based on a dataset of discriminating peaks to successfully type *Mycobacterium abscessus* at the subspecies level. However, their results revealed the need to test a larger, worldwide collection to confirm the hypothesis that different biogeographic MS profiles exist for a given species. In fact, despite the fact that the proposed protocol was independent of the culture method, duration of culture, and the experimenter, none of the specific peaks identified in a previous study [41] were detected when the same extraction protocol was used. An effort to assess the technical and biological reproducibility of MALDI typing between different centres was reported by Oberle et al. [42]. Although their study focused on a small sample size that included only one pathogen, the results revealed that technical and biological reproducibility are sufficient to allow distinguishing clusters to be detected. However, shared standard operating procedures and bioinformatics approaches are required to make the analyses reliable and robust for routine clinical applications. Multicentre validation studies, and proposed good practice guidelines [31, 39] are among the recent joint efforts that will advance the application of proteotyping in clinical laboratories. In addition, construction of dedicated libraries under stringent conditions and the development of software to extract and analyze data could allow MALDI typing to reach the robustness needed to further implement the method in routine clinical laboratory diagnostics. To increase the reliability and the reproducibility of MALDI typing results, Zautner et al. [43] described an approach that combines the analysis of variable masses observed during whole-cell MALDI-TOF mass spectrometry with ribosomal and whole-genome multilocus sequence typing (MLST) database-deduced isoform lists. In contrast to classic whole-spectrum clustering approaches, their strategy considers only changes in mass assigned to one specific set of allelic isoforms for the same protein. The fact that more robust methods rely on peak shifts and integration of genomic/proteomic data was also demonstrated by the successful identification of valuable biomarkers using a proteogenomic approach. In proteogenomics studies, after sequence-based prediction, potential biomarkers are identified by a shotgun proteomics approach before confirming their discriminatory power using linear MALDI-TOF MS profiles [44, 45].

Additional promising results for the implementation of MALDI typing in a clinical setting were obtained by applying methods that exploit the rapidly growing availability of microbial genome sequences to build synthetic reference spectra. Among these methods, the S10-GERMS-based approaches allow proteotyping of bacteria for which genome sequences are not yet decoded, by identifying the nucleotide sequence of genes encoding the ribosomal proteins detected by MALDI-TOF MS, which are specific to bacterial serotypes or strains [46, 47]. So far, this technique has been used to proteotype several clinically-relevant bacteria [48, 49]. As highlighted by the experimental work of Ojima-Kato et al. [49] the construction of databases with the theoretical masses of mass peaks, combined with development of discrimination software, has considerable potential, not only for typing single microorganisms but also for the discrimination, as part of routine diagnostic procedures, of closely-related bacterial strains that are mixed together (e.g., intestinal flora).

The frequent need to deal with polymicrobial samples represents an additional challenge when attempting to routinely type directly from clinical samples. To infer the composition of this type of sample, a few alternative methods have been proposed. The approach described by Mahé et al. [50] relies on a penalized non-negative linear regression framework making use of species-specific prototypes. Within certain limitations, the method automatically determines how many and which species are present in the sample and provides an estimation of their relative concentrations using a single mass spectrum that exploits the same reference database as the one used to identify pure cultures in routine clinical diagnosis. However, the physiological states of the different organisms present in a complex sample may vary considerably, potentially making this approach inappropriate.

Multiple studies have assessed the performance of the commercially-available MALDI-TOF MS systems for routine use (i.e., the MALDI-TOF MS Biotyper (Bruker Daltonics Inc., Germany) and VITEK MS (bioMérieux Inc., France)), and have demonstrated how the accuracy of both systems can be further enhanced by expanding the spectral database [51, 52]. Although both instruments apply similar principles, the procedure and algorithm used to create their individual reference databases are different. As summarized in a recent review by Rahi et al. [52], these differences are reflected in the level of accuracy in the identification of the same set of microorganisms. In addition, although efforts have been made to overcome the limitations of commercial databases by the creation of in-house ones, these libraries tend to be associated with specific instruments and sample preparation methods, limiting the potential for sharing and processing of data produced by different instruments and in different laboratories [53]. The use of a standardized strategy to create error-free and robust databases(s) has been proposed by various authors as a means to improve the efficiency of MALDI identification and typing.

Although most MALDI typing approaches directly use bacteria or whole-cell extracts, further improvements in the discriminatory power of the methodology could be provided by the analysis of specific extracts [31]. A small number of pioneering works [54, 55] have demonstrated how the use of particular bacterial extracts increased the chances of identification of peak biomarkers specific to subgroups. However, the required sample preparation is complex and makes application of this procedure difficult in a routine clinical laboratory, unless automated sample preparation is developed. Recently, to discriminate between clinically-relevant bacterial strains, Fleurbaaij et al. [56] proposed a combination of MALDI-TOF MS-based species identification with ultrahigh resolution 15T MALDI-Fourier transform ion cyclotron resonance (FTICR) MS. The cost of this instrument makes it difficult to envision routine applications. However, more affordable higher resolution instruments could be proposed in the future, and further investigations will be useful in expanding their proof-of-principle experiment. To address the shortcomings of analyzing mass patterns for small, abundant and basic proteins, without actually identifying the proteins, and at the same time to obtain insights into functional traits, Clark et al. [57] recently suggested combining the analysis of mass spectral patterns from conserved housekeeping proteins with signatures for specialized metabolites. However, their innovative pipeline is currently restricted to pure cultures or binary cell mixtures and is limited by the resolution of the instruments used, similar to approaches based on whole-cell MALDI-TOF.

One of the most important limitations of MALDI-TOF MS is its relatively low analytical sensitivity due to its low detection performance for high molecular weight molecules. Depending on the treatment applied to extract proteins,  $10^5$ – $10^7$  bacterial cells are required for efficient MALDI-TOF detection [58]. Thus, identification and typing accuracy is increased when the spectrum is acquired from a colony grown on an agar plate or a culture pellet, e.g. after a culture-based amplification step. However, although it increases accuracy and reduces sample complexity, culturing bacteria from complex mixtures should be expected to favour cultivable microorganisms and could, thereby, limit the number of microorganisms identified.



#### **4. Novel methodology: tandem MS (MS/MS)-based proteotyping**

In the past few years, the potential for successful use of MS-based proteotyping in a clinical setting has been advanced by improvements in the performance of MS instrumentation, especially high-resolution tandem mass spectrometry (MS/MS) combined with chromatography. MS/MS analysis consists of selecting an ion of interest within the first MS for fragmentation into smaller chemical entities, which are then analyzed by the second MS to derive additional structural information. For a peptide, fragmentation by collision with neutral molecules (collision-induced dissociation) produces predictable ions, and the amino acid sequence can then be determined from the  $m/z$  signals recorded in the MS/MS spectrum. The protein/peptide biomarkers are reverse-identified from their sequence-specific fragment ions by comparison with a database of protein sequences. Shotgun proteomics refers to the MS/MS analysis of peptides produced by trypsin proteolysis of all proteins extracted from a sample (Fig. 2). MS/MS and shotgun proteomics provide new potential for the characterization of microorganisms through extensive analysis of their proteins, including taxonomic biomarkers, antibiotic resistance and toxins [2, 11, 59].

Shotgun proteomics offers a higher detection sensitivity and a higher dynamic range compared to the analysis of intact proteins. Today, thousands of peptides from a sample can be resolved and identified during a 60minute MS run performed on the latest high-resolution MS/MS systems. However, one of the main challenges for MS/MS-based proteotyping is the analysis of the data generated. In addition, as highlighted by Karlsson et al. [2], the identification of discriminative peptides that could be used for proteotyping is highly dependent on access to a comprehensive and accurately curated database. Various groups have prioritized the development of bioinformatics tools based on robust algorithms to classify informative peptide fragments and identify and characterize organisms without prior information. For example, Dworzanski et al. [60] proposed a proteomics approach based on the use of fractions of shared peptides as an estimator of relatedness among closely-related species. A database of 170 fully-sequenced bacterial genomes was used to create a matrix of sequence-to-bacterium assignments and to interpret spectra from MS/MS experiments. Following this approach, serotypes for *B. cereus* strains could be successfully distinguished [61] without requiring whole-genome sequencing to type strains. As closely-related organisms share some peptide sequences, this method is, in principle, applicable to non-genome-sequenced bacteria. However, limitations do exist (e.g., need for pure cultures) that compromise its clinical applicability. A new computational method (TCUP) was recently presented by Boulund et al. [62]. By comparing the protein-sequence data generated to references databases, TCUP automatically selects peptides suitable for characterization of taxonomic composition and identification of expressed antimicrobial resistance genes. The method was also shown capable of identifying and estimating the relative abundance of individual species in mixed samples, thereby confirming its potential to further extend the use of bottom-up MS/MS in clinical settings. An important drawback of these methods is linked to composition of the databases. Searching comprehensive databases in order to maximize identification is not only computationally inefficient, but also potentially increases the number of false positive identifications. Moreover, the increasing number of sequenced genomes makes it more challenging to identify specific peptides at species or subspecies levels, since an increased number of peptides will be shared by closely-related microorganisms.

Other alternatives based on MS/MS are possible, as exemplified by the method based on unidentified tandem mass spectral libraries proposed by Önder et al. [63], a workflow which is peptide identification-free, and thus genome sequence-independent. This method uses a similarity cluster algorithm to group MS/MS spectra that may be derived from the same peptide ion and,

subsequently, merges them into a unique consensus spectrum to generate MS/MS signal fingerprints. Shao et al. [64] used this workflow to successfully differentiate bacteria at the strain level. However, although the fingerprints generated are richer in information and, hence, potentially more discriminative when differentiating bacterial isolates, the reliability of the method depends on the effectiveness of the filtering algorithm to detect and eliminate MS/MS spectra derived from non-peptides, as well as on the sensitivity and specificity of the clustering algorithm used to distinguish spectra from one peptide versus another. Furthermore, this approach has not yet been shown to work on complex samples containing at least ten distinct microorganisms.

Targeted approaches based on the characterization of specific sets of peptides by selected reaction monitoring mode (SRM) are among the methods based on MS/MS that have proven to be particularly useful for proteotyping and, more generally, for clinical applications. Characterized by a high multiplexing capacity, SRM allows the development of assays to quantify proteoform of a given protein in a sample, provided the different forms are characterized by a mass shift. In SRM methods, pre-selected peptides are quantified by monitoring specific precursor-to-product ion transitions using MS/MS [65]. Charretier et al. [66] demonstrated the high multiplexing capabilities of the approach by using the most specific peptides to design an SRM method for in-depth characterization of *S. aureus* strains in 60–80 minutes using a single, multiplexed analysis. The method not only correctly determined the antibiotic resistance of the strains, it also identified two relevant toxins. SRM illustrates the potential clinical applicability of MS/MS in developing personalized patient treatments, particularly in the context of the emergence of multidrug resistance and of the current dearth of therapeutic options. Notwithstanding, potential pitfalls like the occurrence of false positives due to sequence errors, or noise and interference from the sample matrix, or the existence of mixed cultures among clinical samples, must be addressed in future developments of targeted approaches if we are to see wider clinical application.

## **5. Limitations**

In terms of instrumentation, successful typing requires high resolution, high mass accuracy, high sensitivity, a wide dynamic range and high scan rate [2]. MALDI-TOF MS-based typing methods are simple and high throughput, but the quality of the MS profile mass accuracy, data richness and reproducibility could be further improved to attain the discriminative power required to reliably differentiate between microorganisms at the subspecies level, as required for epidemiological surveys. The indubitable advantages of implementing MS/MS-based proteotyping approaches in routine clinical practice include their theoretically greater discriminatory capability than methods that detect small intact proteins. However, several improvements are still required to overcome limitations of the currently available instrumentation and approaches. First, MS/MS-based methods are much more demanding in terms of time, effort and expertise required for sample preparation, analysis and data processing. These aspects currently represent major constraints preventing the implementation of MS/MS-based proteotyping in routine diagnostics. In this context, development of user-friendly software to facilitate handling of the acquired data is a priority. Second, the cost of high-resolution instruments may restrict their adoption and use. Furthermore, a technical limitation of MS/MS-based proteotyping methods is the requirement for a denser and more accurate reference database. In fact, to be successful, proteotyping requires comprehensive and accurate whole-genome sequence databases and a stable taxonomy. The results produced will only be as accurate as the databases available, which becomes problematic when we consider the high rates of misclassified WGS data observed in public databases [67, 68], or when dealing with new species or emerging pathogens, for which data might not be as readily available as for other species. Last but

not least, all of the MS/MS methods cited require further simplification, automation, and validation before they can be implemented as part of a clinical routine. Specifically, if we are to overcome the additional challenge represented by the presence of human biomass such as cells, mucus, and proteins, sample preparation must be optimized to allow direct typing of pathogens from clinical specimens without culturing.

## **6. Perspectives**

MALDI-TOF MS proteotyping protocols have been extended and improved in many respects over the last decade, and this technology has become a versatile tool that can be useful beyond the identification of a bacterial species, e.g. for the characterization of arthropod vectors of infectious diseases including ticks and mosquitoes [69].

Optimized typing strategies (e.g., control and standardization of culture conditions, sample preparation procedures, and sample analysis methods) are needed to enhance the reproducibility and portability of proteomics data. To reduce the time-to-answer, approaches are increasingly being designed to directly process a range of medical specimens (e.g., blood and urine samples) [70]. Differential centrifugation, where blood cells are first removed from the sample by low-speed centrifugation, was one of the first techniques proposed to analyze microbial mixtures in blood cultures [71]. Differential lysis-based methods, involving lysis of blood cells from blood cultures prior to extraction of bacterial proteins, have been reported to be faster, simpler, and to offer better performance [72]. Inigo et al. [73] proposed a combining a urine screening method, such as flow cytometry or automated microscopic urine sediment analysis, with MALDI-TOF MS for reliable identification of bacterial strains, especially Gram-negative, from urine samples,

The differentiation of bacterial, viral and other pathogenic isolates represents only one of the promising and clinically-relevant applications of proteotyping. Innovative proteotyping approaches are currently being developed to gain insight into host defence responses, to establish potentially useful targets for therapeutic intervention, and to identify biomarkers to identify the affected organs during a bacterial infection. Along these lines, a promising approach was recently proposed by Lapek et al. [74] that demonstrated the potential of their proteomics-based strategy for the identification of organ-specific and plasma-trackable markers of Group A *Streptococcus* (GAS) infection. Advanced, culture independent, MS-based typing approaches could represent a valuable tool to assess the structure of the microbiome at a specific body site. By typing its components, proteotyping could represent a new diagnostic and disease-monitoring strategy that contributes to evaluation of the extent of dysbiosis or the co-occurrence of pathogens. In this context, microbiome proteotyping could aid medical decision-making regarding selection of antimicrobial agents and management of the related disease.

The importance of proteotypes in deciphering key biological insights has also been pointed out in the field of cancer biology and, more generally, the characterization of mammalian cells [75]. Guo et al. [76] described a rapid proteotyping approach to reveal cancer biology determinants and identified novel drug-response determinants for clinically-relevant chemotherapeutic and targeted therapies. In addition, Munteanu and Hopf [77] discussed the adoption of MS-based workflows for the classification of mammalian cells highlighting emerging applications for this methodology in clinical diagnostics, among other areas. However, in this context the potential of proteotyping and its translational clinical applications has yet to be fully exploited.

## **7. Concluding remarks**

Over the last decade, MALDI-TOF MS proteotyping became a standard method in clinical microbiology laboratories for routine identification of bacterial pathogens. As discussed in this review, this approach will continue to be improved, increasing in accuracy and sensitivity. Recent developments suggest that clinical proteotyping will experience another dramatic revolution. MS/MS of peptides demonstrates excellent sensitivity compared to MALDI-TOF of small intact proteins, and, in addition, allows many more discriminative signals to be recorded. This method will probably be used in two complementary ways; first, targeted proteomics could be further developed to identify the most relevant bacterial groups and their associated antibiotic resistances; alternatively, shotgun proteomics could appear soon as a powerful method to identify any kind of bacteria present in medical samples, without the need for prior knowledge, while also providing a quick antibiotic-resistance checklist based on the detected proteins. This final proteotyping method could be extremely valuable when analyzing complex samples such as the human microbiota.

### **Acknowledgements**

We thank the “Commissariat à l’Energie Atomique et aux Energies Alternatives” (France), and the ANR programme “Phylopeptidomics” [grant number ANR-17-CE18-0023] for financial support. We are indebted to all our close colleagues, microbiologists, clinicians, mass spectrometrists, proteomists, and bioinformaticians for sharing their enthusiasm for proteotyping.

### **Conflict of Interest**

None of the authors has any conflicts of interest to disclose.

### **Figure Legends**

**Figure 1.** Advantages of implementing proteotyping in the routine clinical practice.

**Figure 2.** Current and novel methodologies for mass spectrometry-based pathogen typing. Relevant references are indicated in squared brackets.

## References

1. Sloan A, Wang G, Cheng K: **Traditional approaches versus mass spectrometry in bacterial identification and typing.** *Clin Chim Acta* 2017, **473**:180-185.
2. Karlsson R, Gonzales-Siles L, Boulund F, Svensson-Stadler L, Skovbjerg S, Karlsson A, Davidson M, Hulth S, Kristiansson E, Moore ER: **Proteotyping: Proteomic characterization, classification and identification of microorganisms--A prospectus.** *Syst Appl Microbiol* 2015, **38**(4):246-257.
3. Wilmes P, Heintz-Buschart A, Bond PL: **A decade of metaproteomics: where we stand and what the future holds.** *Proteomics* 2015, **15**(20):3409-3417.
4. van Belkum A, Welker M, Pincus D, Charrier JP, Girard V: **Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry in Clinical Microbiology: What Are the Current Issues?** *Ann Lab Med* 2017, **37**(6):475-483.
5. Lavigne JP, Espinal P, Dunyach-Remy C, Messad N, Pantel A, Sotto A: **Mass spectrometry: a revolution in clinical microbiology?** *Clin Chem Lab Med* 2013, **51**(2):257-270.
6. Mlaga KD, Dubourg G, Abat C, Chaudet H, Lotte L, Diene SM, Raoult D, Ruimy R, Rolain JM: **Using MALDI-TOF MS typing method to decipher outbreak: the case of Staphylococcus saprophyticus causing urinary tract infections (UTIs) in Marseille, France.** *Eur J Clin Microbiol Infect Dis* 2017, **36**(12):2371-2377.
7. Singhal N, Kumar M, Kanaujia PK, Virdi JS: **MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis.** *Front Microbiol* 2015, **6**:791.
8. Qian J, Cutler JE, Cole RB, Cai Y: **MALDI-TOF mass signatures for differentiation of yeast species, strain grouping and monitoring of morphogenesis markers.** *Anal Bioanal Chem* 2008, **392**(3):439-449.
9. Pulcrano G, Roscetto E, Iula VD, Panellis D, Rossano F, Catania MR: **MALDI-TOF mass spectrometry and microsatellite markers to evaluate Candida parapsilosis transmission in neonatal intensive care units.** *Eur J Clin Microbiol Infect Dis* 2012, **31**(11):2919-2928.
10. Santos CRF, E.; Mazza, M.; Padovan, A. C. B.; Colombo, A.; Lima, N.: **Impact of MALDI-TOF MS in Clinical Mycology; Progress and Barriers in Diagnostics.** In: *MALDI-TOF and Tandem MS for Clinical Microbiology.* Edited by Shah HNG, S.E.; 2017.
11. Cheng K, Chui H, Domish L, Hernandez D, Wang G: **Recent development of mass spectrometry and proteomics applications in identification and typing of bacteria.** *Proteomics Clin Appl* 2016, **10**(4):346-357.
12. Maurer FP, Christner M, Hentschke M, Rohde H: **Advances in Rapid Identification and Susceptibility Testing of Bacteria in the Clinical Microbiology Laboratory: Implications for Patient Care and Antimicrobial Stewardship Programs.** *Infect Dis Rep* 2017, **9**(1):6839.
13. Vrioni G, Tsiamis C, Oikonomidis G, Theodoridou K, Kapsimali V, Tsakris A: **MALDI-TOF mass spectrometry technology for detecting biomarkers of antimicrobial resistance: current achievements and future perspectives.** *Ann Transl Med* 2018, **6**(12):240.
14. Hrabak J, Chudackova E, Walkova R: **Matrix-assisted laser desorption ionization-time of flight (maldi-tof) mass spectrometry for detection of antibiotic resistance mechanisms: from research to routine diagnosis.** *Clin Microbiol Rev* 2013, **26**(1):103-114.
15. Perez-Llarena FJ, Bou G: **Proteomics As a Tool for Studying Bacterial Virulence and Antimicrobial Resistance.** *Front Microbiol* 2016, **7**:410.
16. Kostrzewa M, Pranada AB: **Future Applications of MALDI-TOF MS in Microbiology.** In: *Advances in MALDI and Laser-Induced Soft Ionization Mass Spectrometry.* Edited by Cramer R. Cham: Springer International Publishing; 2016: 227-246.
17. Josten M, Dischinger J, Szekat C, Reif M, Al-Sabti N, Sahl HG, Parcina M, Bekeredjian-Ding I, Bierbaum G: **Identification of agr-positive methicillin-resistant Staphylococcus aureus harbouring the class A mec complex by MALDI-TOF mass spectrometry.** *Int J Med Microbiol* 2014, **304**(8):1018-1023.

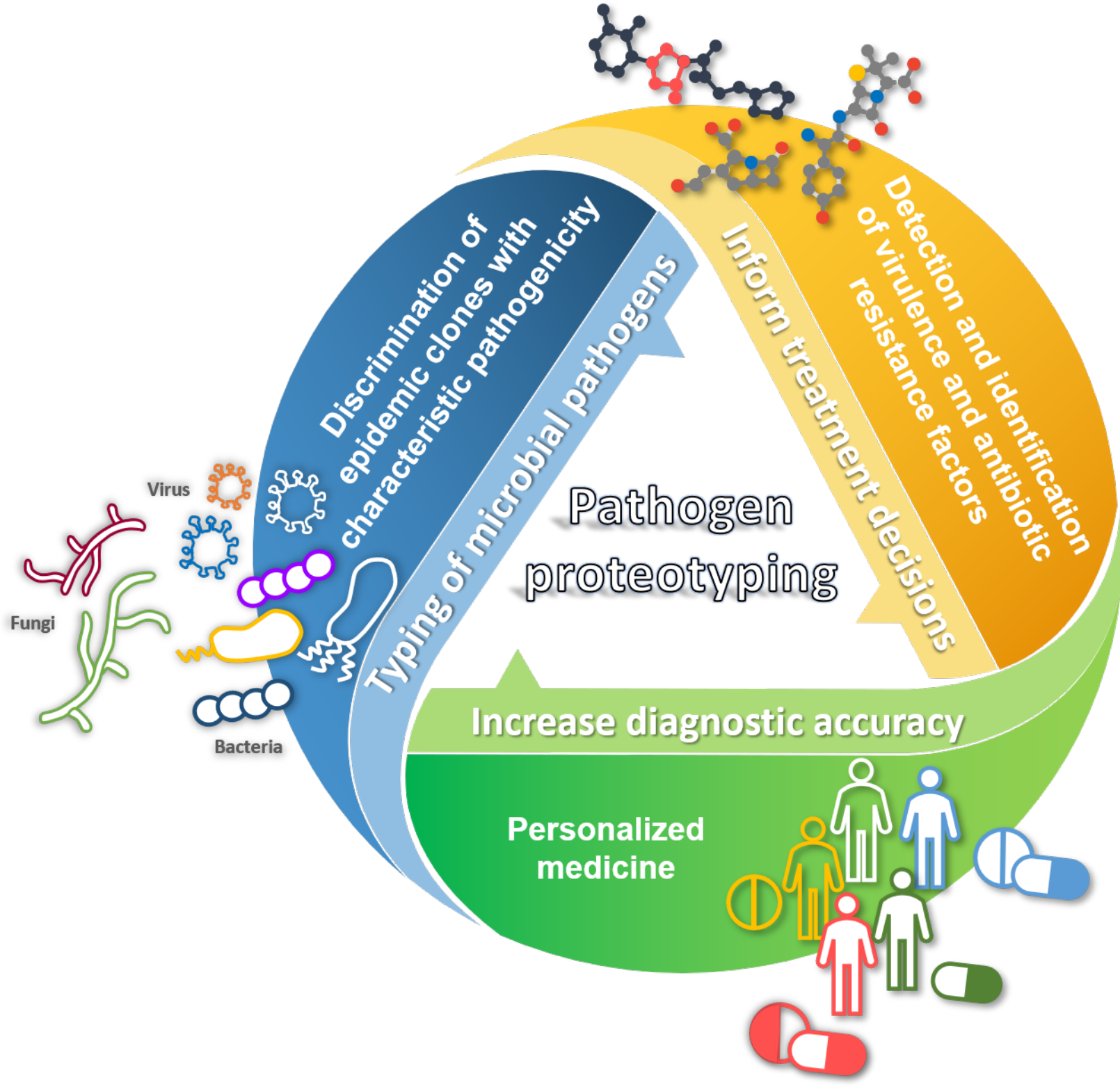
18. Nagy E, Becker S, Soki J, Urban E, Kostrzewa M: **Differentiation of division I (cfiA-negative) and division II (cfiA-positive) Bacteroides fragilis strains by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.** *J Med Microbiol* 2011, **60**(Pt 11):1584-1590.
19. Fagerquist CK, Zaragoza WJ, Sultan O, Woo N, Quinones B, Cooley MB, Mandrell RE: **Top-down proteomic identification of Shiga toxin 2 subtypes from Shiga toxin-producing Escherichia coli by matrix-assisted laser desorption ionization-tandem time of flight mass spectrometry.** *Appl Environ Microbiol* 2014, **80**(9):2928-2940.
20. Bizzini A, Greub G: **Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification.** *Clin Microbiol Infect* 2010, **16**(11):1614-1619.
21. Croxatto A, Prod'hom G, Greub G: **Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology.** *FEMS Microbiol Rev* 2012, **36**(2):380-407.
22. Welker M: **Proteomics for routine identification of microorganisms.** *Proteomics* 2011, **11**(15):3143-3153.
23. Sandrin TR, Goldstein JE, Schumaker S: **MALDI TOF MS profiling of bacteria at the strain level: a review.** *Mass Spectrom Rev* 2013, **32**(3):188-217.
24. Karger A: **Current developments to use linear MALDI-TOF spectra for the identification and typing of bacteria and the characterization of other cells/organisms related to infectious diseases.** *Proteomics Clin Appl* 2016, **10**(9-10):982-993.
25. Durighello E, Bellanger L, Ezan E, Armengaud J: **Proteogenomic biomarkers for identification of Francisella species and subspecies by matrix-assisted laser desorption ionization-time-of-flight mass spectrometry.** *Anal Chem* 2014, **86**(19):9394-9398.
26. Mesureur J, Arend S, Celliere B, Courault P, Cotte-Pattat PJ, Totty H, Deol P, Mick V, Girard V, Touchberry J *et al*: **A MALDI-TOF MS database with broad genus coverage for species-level identification of Brucella.** *PLoS Negl Trop Dis* 2018, **12**(10):e0006874.
27. Ueda O, Tanaka S, Nagasawa Z, Hanaki H, Shobuike T, Miyamoto H: **Development of a novel matrix-assisted laser desorption/ionization time-of-flight mass spectrum (MALDI-TOF-MS)-based typing method to identify meticillin-resistant Staphylococcus aureus clones.** *J Hosp Infect* 2015, **90**(2):147-155.
28. Purighalla S, Esakimuthu S, Reddy M, Varghese GK, Richard VS, Sambandamurthy VK: **Discriminatory power of three typing techniques in determining relatedness of nosocomial Klebsiella pneumoniae isolates from a tertiary hospital in India.** *Indian J Med Microbiol* 2017, **35**(3):361-368.
29. Pinto TC, Costa NS, Castro LF, Ribeiro RL, Botelho AC, Neves FP, Peralta JM, Teixeira LM: **Potential of MALDI-TOF MS as an alternative approach for capsular typing Streptococcus pneumoniae isolates.** *Sci Rep* 2017, **7**:45572.
30. Murray PR: **Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: usefulness for taxonomy and epidemiology.** *Clin Microbiol Infect* 2010, **16**(11):1626-1630.
31. Sauguet M, Valot B, Bertrand X, Hocquet D: **Can MALDI-TOF Mass Spectrometry Reasonably Type Bacteria?** *Trends Microbiol* 2017, **25**(6):447-455.
32. Ghyselincx J, Van Hoorde K, Hoste B, Heylen K, De Vos P: **Evaluation of MALDI-TOF MS as a tool for high-throughput dereplication.** *J Microbiol Methods* 2011, **86**(3):327-336.
33. Marin M, Cercenado E, Sanchez-Carrillo C, Ruiz A, Gomez Gonzalez A, Rodriguez-Sanchez B, Bouza E: **Accurate Differentiation of Streptococcus pneumoniae from other Species within the Streptococcus mitis Group by Peak Analysis Using MALDI-TOF MS.** *Front Microbiol* 2017, **8**:698.
34. Rodrigues NM, Bronzato GF, Santiago GS, Botelho LA, Moreira BM, Coelho ID, Souza MM, Coelho SM: **The Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) identification versus biochemical tests: a study with enterobacteria from a dairy cattle environment.** *Braz J Microbiol* 2017, **48**(1):132-138.

35. Gekenidis MT, Studer P, Wuthrich S, Brunisholz R, Drissner D: **Beyond the matrix-assisted laser desorption ionization (MALDI) biotyping workflow: in search of microorganism-specific tryptic peptides enabling discrimination of subspecies.** *Appl Environ Microbiol* 2014, **80**(14):4234-4241.
36. Madonna AJ, Basile F, Ferrer I, Meetani MA, Rees JC, Voorhees KJ: **On-probe sample pretreatment for the detection of proteins above 15 KDa from whole cell bacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.** *Rapid Commun Mass Spectrom* 2000, **14**(23):2220-2229.
37. Dieckmann R, Helmuth R, Erhard M, Malorny B: **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* 2008, **74**(24):7767-7778.
38. Dieckmann R, Malorny B: **Rapid screening of epidemiologically important Salmonella enterica subsp. enterica serovars by whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry.** *Appl Environ Microbiol* 2011, **77**(12):4136-4146.
39. Spinali S, van Belkum A, Goering RV, Girard V, Welker M, Van Nuenen M, Pincus DH, Arzac M, Durand G: **Microbial typing by matrix-assisted laser desorption ionization-time of flight mass spectrometry: do we need guidance for data interpretation?** *J Clin Microbiol* 2015, **53**(3):760-765.
40. Fangous MS, Mougari F, Gouriou S, Calvez E, Raskine L, Cambau E, Payan C, Hery-Arnaud G: **Classification algorithm for subspecies identification within the Mycobacterium abscessus species, based on matrix-assisted laser desorption ionization-time of flight mass spectrometry.** *J Clin Microbiol* 2014, **52**(9):3362-3369.
41. Teng SH, Chen CM, Lee MR, Lee TF, Chien KY, Teng LJ, Hsueh PR: **Matrix-assisted laser desorption ionization-time of flight mass spectrometry can accurately differentiate between Mycobacterium masillense (M. abscessus subspecies bolletti) and M. abscessus (sensu stricto).** *J Clin Microbiol* 2013, **51**(9):3113-3116.
42. Oberle M, Wohlwend N, Jonas D, Maurer FP, Jost G, Tschudin-Sutter S, Vranckx K, Egli A: **The Technical and Biological Reproducibility of Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) Based Typing: Employment of Bioinformatics in a Multicenter Study.** *PLoS One* 2016, **11**(10):e0164260.
43. Zautner AE, Masanta WO, Weig M, Gross U, Bader O: **Mass Spectrometry-based PhyloProteomics (MSPP): A novel microbial typing Method.** *Sci Rep* 2015, **5**:13431.
44. Suarez S, Ferroni A, Lotz A, Jolley KA, Guerin P, Leto J, Dauphin B, Jamet A, Maiden MC, Nassif X *et al*: **Ribosomal proteins as biomarkers for bacterial identification by mass spectrometry in the clinical microbiology laboratory.** *J Microbiol Methods* 2013, **94**(3):390-396.
45. Armengaud J: **Defining Diagnostic Biomarkers Using Shotgun Proteomics and MALDI-TOF Mass Spectrometry.** *Methods Mol Biol* 2017, **1616**:107-120.
46. Hotta Y, Teramoto K, Sato H, Yoshikawa H, Hosoda A, Tamura H: **Classification of genus Pseudomonas by MALDI-TOF MS based on ribosomal protein coding in S10-spc-alpha operon at strain level.** *J Proteome Res* 2010, **9**(12):6722-6728.
47. Ojima-Kato T, Yamamoto N, Nagai S, Shima K, Akiyama Y, Ota J, Tamura H: **Application of proteotyping Strain Solution ver. 2 software and theoretically calculated mass database in MALDI-TOF MS typing of Salmonella serotype.** *Appl Microbiol Biotechnol* 2017, **101**(23-24):8557-8569.
48. Hotta Y, Sato J, Sato H, Hosoda A, Tamura H: **Classification of the genus Bacillus based on MALDI-TOF MS analysis of ribosomal proteins coded in S10 and spc operons.** *J Agric Food Chem* 2011, **59**(10):5222-5230.
49. Ojima-Kato T, Yamamoto N, Iijima Y, Tamura H: **Assessing the performance of novel software Strain Solution on automated discrimination of Escherichia coli serotypes and their mixtures using matrix-assisted laser desorption ionization-time of flight mass spectrometry.** *J Microbiol Methods* 2015, **119**:233-238.

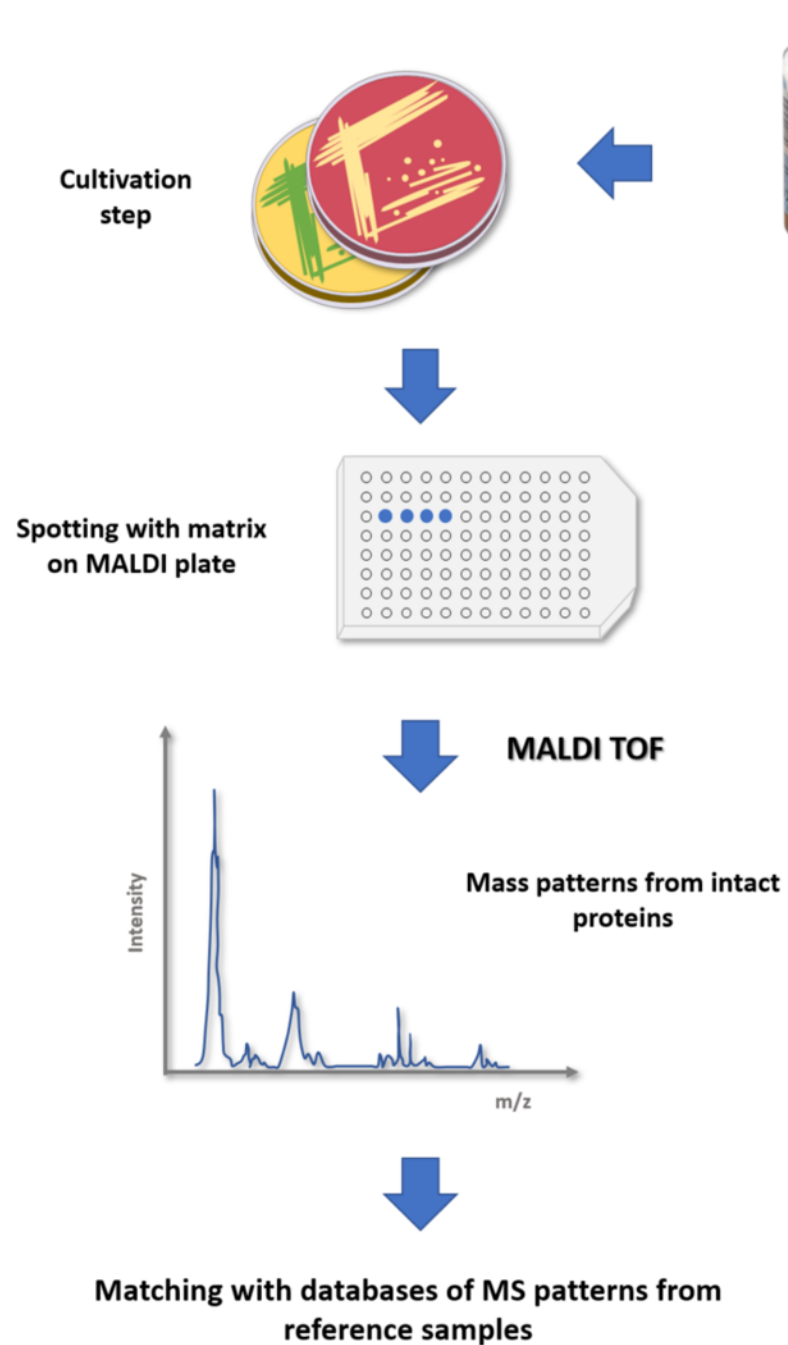
50. Mahe P, Arsac M, Chatellier S, Monnin V, Perrot N, Mailler S, Girard V, Ramjeet M, Surre J, Lacroix B *et al*: **Automatic identification of mixed bacterial species fingerprints in a MALDI-TOF mass-spectrum.** *Bioinformatics* 2014, **30**(9):1280-1286.
51. Deak E, Charlton CL, Bobenchik AM, Miller SA, Pollett S, McHardy IH, Wu MT, Garner OB: **Comparison of the Vitek MS and Bruker Microflex LT MALDI-TOF MS platforms for routine identification of commonly isolated bacteria and yeast in the clinical microbiology laboratory.** *Diagn Microbiol Infect Dis* 2015, **81**(1):27-33.
52. Rahi P, Prakash O, Shouche YS: **Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass-Spectrometry (MALDI-TOF MS) Based Microbial Identifications: Challenges and Scopes for Microbial Ecologists.** *Front Microbiol* 2016, **7**:1359.
53. Cheng D, Qiao L, Horvatovich P: **Toward Spectral Library-Free Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Bacterial Identification.** *J Proteome Res* 2018, **17**(6):2124-2130.
54. Chui H, Chan M, Hernandez D, Chong P, McCorrister S, Robinson A, Walker M, Peterson LA, Ratnam S, Haldane DJ *et al*: **Rapid, Sensitive, and Specific Escherichia coli H Antigen Typing by Matrix-Assisted Laser Desorption Ionization-Time of Flight-Based Peptide Mass Fingerprinting.** *J Clin Microbiol* 2015, **53**(8):2480-2485.
55. Cox CR, Jensen KR, Saichek NR, Voorhees KJ: **Strain-level bacterial identification by CeO<sub>2</sub>-catalyzed MALDI-TOF MS fatty acid analysis and comparison to commercial protein-based methods.** *Sci Rep* 2015, **5**:10470.
56. Fleurbaaij F, Kraakman ME, Claas EC, Knetsch CW, van Leeuwen HC, van der Burgt YE, Veldkamp KE, Vos MC, Goessens W, Mertens BJ *et al*: **Typing Pseudomonas aeruginosa Isolates with Ultrahigh Resolution MALDI-FTICR Mass Spectrometry.** *Anal Chem* 2016, **88**(11):5996-6003.
57. Clark CM, Costa MS, Sanchez LM, Murphy BT: **Coupling MALDI-TOF mass spectrometry protein and specialized metabolite analyses to rapidly discriminate bacterial function.** *Proc Natl Acad Sci U S A* 2018, **115**(19):4981-4986.
58. Carbonnelle E, Mesquita C, Bille E, Day N, Dauphin B, Beretti JL, Ferroni A, Gutmann L, Nassif X: **MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory.** *Clin Biochem* 2011, **44**(1):104-109.
59. Armengaud J: **Microbiology and proteomics, getting the best of both worlds!** *Environ Microbiol* 2013, **15**(1):12-23.
60. Dworzanski JP, Deshpande SV, Chen R, Jabbour RE, Snyder AP, Wick CH, Li L: **Mass spectrometry-based proteomics combined with bioinformatic tools for bacterial classification.** *J Proteome Res* 2006, **5**(1):76-87.
61. Dworzanski JP, Dickinson DN, Deshpande SV, Snyder AP, Eckenrode BA: **Discrimination and phylogenomic classification of Bacillus anthracis-cereus-thuringiensis strains based on LC-MS/MS analysis of whole cell protein digests.** *Anal Chem* 2010, **82**(1):145-155.
62. Boulund F, Karlsson R, Gonzales-Siles L, Johnning A, Karami N, Al-Bayati O, Ahren C, Moore ERB, Kristiansson E: **Typing and Characterization of Bacteria Using Bottom-up Tandem Mass Spectrometry Proteomics.** *Mol Cell Proteomics* 2017, **16**(6):1052-1063.
63. Onder O, Shao W, Kemps BD, Lam H, Brisson D: **Identifying sources of tick blood meals using unidentified tandem mass spectral libraries.** *Nat Commun* 2013, **4**:1746.
64. Shao W, Zhang M, Lam H, Lau SC: **A peptide identification-free, genome sequence-independent shotgun proteomics workflow for strain-level bacterial differentiation.** *Sci Rep* 2015, **5**:14337.
65. Picotti P, Bodenmiller B, Aebersold R: **Proteomics meets the scientific method.** *Nat Methods* 2013, **10**(1):24-27.
66. Charretier Y, Dauwalder O, Franceschi C, Degout-Charmette E, Zambardi G, Cecchini T, Bardet C, Lacoux X, Dufour P, Veron L *et al*: **Rapid Bacterial Identification, Resistance, Virulence and Type Profiling using Selected Reaction Monitoring Mass Spectrometry.** *Sci Rep* 2015, **5**:13944.



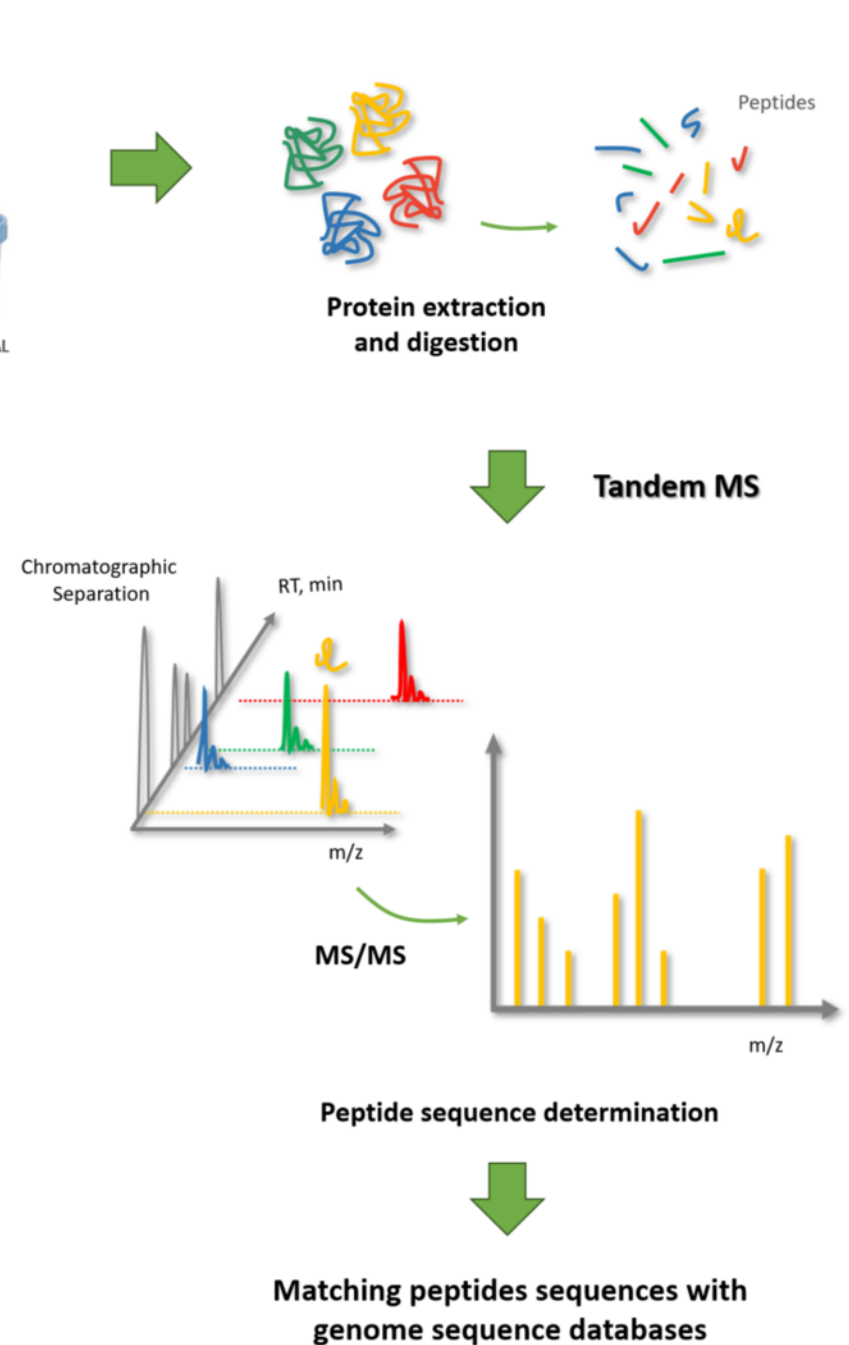
67. Pible O, Armengaud J: **Improving the quality of genome, protein sequence, and taxonomy databases: a prerequisite for microbiome meta-omics 2.0.** *Proteomics* 2015, **15**(20):3418-3423.
68. Ciufu S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M: **Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI.** *Int J Syst Evol Microbiol* 2018, **68**(7):2386-2392.
69. Yssouf A, Almeras L, Raoult D, Parola P: **Emerging tools for identification of arthropod vectors.** *Future Microbiol* 2016, **11**(4):549-566.
70. Sandrin TR, Demirev PA: **Characterization of microbial mixtures by mass spectrometry.** *Mass Spectrom Rev* 2018, **37**(3):321-349.
71. Christner M, Rohde H, Wolters M, Sobottka I, Wegscheider K, Aepfelbacher M: **Rapid identification of bacteria from positive blood culture bottles by use of matrix-assisted laser desorption-ionization time of flight mass spectrometry fingerprinting.** *J Clin Microbiol* 2010, **48**(5):1584-1591.
72. Nonnemann B, Tvede M, Bjarnsholt T: **Identification of pathogenic microorganisms directly from positive blood vials by matrix-assisted laser desorption/ionization time of flight mass spectrometry.** *APMIS* 2013, **121**(9):871-877.
73. Inigo M, Coello A, Fernandez-Rivas G, Rivaya B, Hidalgo J, Quesada MD, Ausina V: **Direct Identification of Urinary Tract Pathogens from Urine Samples, Combining Urine Screening Methods and Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry.** *J Clin Microbiol* 2016, **54**(4):988-993.
74. Lapek JD, Jr., Mills RH, Wozniak JM, Campeau A, Fang RH, Wei X, van de Groep K, Perez-Lopez A, van Sorge NM, Raffatellu M *et al*: **Defining Host Responses during Systemic Bacterial Infection through Construction of a Murine Organ Proteome Atlas.** *Cell Syst* 2018, **6**(5):579-592 e574.
75. Povey JF, O'Malley CJ, Root T, Martin EB, Montague GA, Feary M, Trim C, Lang DA, Alldread R, Racher AJ *et al*: **Rapid high-throughput characterisation, classification and selection of recombinant mammalian cell line phenotypes using intact cell MALDI-ToF mass spectrometry fingerprinting and PLS-DA modelling.** *J Biotechnol* 2014, **184**:84-93.
76. Guo T, Luna A, Koh CC, Rajapakse V, Wu Z, Menden MP, Cheng Y, Calzone L, Martignetti L, Ori A *et al*: **Rapid proteotyping reveals cancer biology and drug response determinants in the NCI-60 cells.** 2018.
77. Munteanu B, Hopf C: **Emergence of whole-cell MALDI-MS biotyping for high-throughput bioanalysis of mammalian cells?** *Bioanalysis* 2013, **5**(8):885-893.



## MALDI –based typing



## Tandem MS –based typing



### Advantages

Speed, Low-cost, Simplicity, Suitable for typing a wide range of microbes

### Recent improvements

- Improvement of spectral databases [26-28] and construction of synthetic reference spectra [49-50]
- Use of broader mass range (from 15 to 75 kDa) [36-38]
- Development of proteogenomics approaches [44-46]
- Use of classification algorithms based on data set of discriminating peaks [40]
- Multicentre validation studies and use of good practice guidelines [31, 39, 41]

Higher detection sensitivity and higher dynamic range than MALDI –based typing, Suitable for the analysis of polymicrobial samples

- Development of approach based on the use of fractions of shared peptides as estimator of the relatedness among closely related species [61]
- Development of computational method (TCUP) [63]
- Targeted approaches based on the characterization of specific sets of peptides by selected reaction monitoring mode (SRM) [66-67]