

# La spectrométrie de masse MALDI-TOF : un outil de diagnostic pour l'identification et le typage des bactéries pathogènes des poissons

Eric DUCHAUD<sup>1</sup>

## CORRESPONDANCE

[eric.duchaud@inrae.fr](mailto:eric.duchaud@inrae.fr)

## RÉSUMÉ

Les méthodes d'identification des bactéries responsables d'épisodes infectieux chez les poissons reposent encore essentiellement sur des techniques traditionnelles, pas toujours discriminantes, comme l'histologie, la cytologie, la recherche de caractères phénotypiques et le sérotypage. Des méthodes moléculaires modernes (e.g., MLST, WGS) tendent à se substituer à ces approches classiques, mais restent le plus souvent, de par leur coût, leur débit et le temps requis, utilisées pour des études rétrospectives dans le cadre de travaux académiques. Pour adapter les traitements et prendre les mesures visant à contrôler un épisode infectieux, les aquaculteurs et les différentes parties prenantes (e.g., vétérinaires, Groupements de Défense Sanitaire Aquacole) ont besoin d'un diagnostic précis, rapide et, idéalement, peu onéreux de l'agent causal. La spectrométrie de masse à temps de vol par désorption laser assistée par matrice (MALDI-TOF MS) est une technique d'identification bactérienne largement répandue dans les laboratoires d'analyses de microbiologie médicale, qui tend à s'étendre aux laboratoires de diagnostic vétérinaire. Basée sur l'analyse directe de cultures bactériennes, la technique MALDI-TOF MS génère un spectre protéique qui est ensuite comparé à un ensemble de spectres de référence stockés dans une base de données. C'est une technique rapide, bon marché et qui se prête bien aux analyses à haut-débit. Cependant, les bases de données commerciales MALDI-TOF MS ne contiennent que peu d'espèces de bactéries pathogènes des poissons. D'autre part, l'identification est insuffisante pour répondre à des questions relevant de l'épidémiologie, dont la base est la caractérisation fine (typage) des isolats bactériens. Nous avons développé un schéma MALDI-TOF MS d'identification pour les espèces du genre *Tenacibaculum* et amélioré ses performances pour son utilisation comme outil épidémiologique de première ligne pour l'espèce *T. maritimum*, un redoutable agent pathogène des poissons marins. Cet article intègre des conseils issus de cette expérience et de recherches bibliographiques, afin de guider le lecteur pour le développement d'une telle approche.

## MOTS-CLÉS

Diagnostic, bactéries pathogènes des poissons, identification, typage, MALDI-TOF MS.

<sup>1</sup> Unité Virologie et Immunologie Moléculaires, INRAE, Université Paris-Saclay, 78350 Jouy-en-Josas, France.

# MALDI-TOF mass spectrometry: a diagnostic tool for identifying and typing fish pathogenic bacteria

Eric DUCHAUD<sup>1</sup>

## CORRESPONDENCE

[eric.duchaud@inrae.fr](mailto:eric.duchaud@inrae.fr)

## ABSTRACT

Methods used to identify the bacteria responsible for outbreaks of infectious diseases in fish farms still essentially rely on traditional and usually non-discriminatory techniques, such as histology, cytology, phenotypic characteristics and serotyping. Modern molecular methods (e.g., MLST, WGS) tend to replace these classical approaches but are usually employed for retrospective studies within the framework of academic research due to their cost, throughput and the time required. To adapt treatments and take control measures, fish farmers and stakeholders (e.g., aquaculture veterinarians) need an accurate, rapid and, ideally, inexpensive diagnosis of the causative agent. Matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF MS) is a bacterial identification technique widely used in medical microbiology laboratories that is progressively spreading to veterinary diagnostic laboratories. Based on the direct analysis of bacterial cultures, MALDI-TOF MS generates a protein spectrum which is then compared to a set of reference spectra stored in a database. It is a fast, inexpensive technique well-adapted to high-throughput analyses. However, commercial MALDI-TOF MS databases contain only a limited number of bacterial fish pathogens. On the other hand, identification is insufficient to answer questions related to epidemiology, the basis of which is the fine characterization (typing) of bacterial isolates. We have developed a MALDI-TOF MS identification method for species of the genus *Tenacibaculum* and improved its performance for use as a first-line epidemiological tool for the species *T. maritimum*, a devastating pathogen of marine fish. This technical notebook includes advice resulting from this experiment and bibliographic research to guide the reader in developing such an approach.

## KEYWORDS

Diagnostics, fish bacterial pathogens, identification, typing, MALDI-TOF MS.

---

<sup>1</sup> Unité Virologie et Immunologie Moléculaires, INRAE, Université Paris-Saclay, 78350 Jouy-en-Josas, France.

## Introduction

L'aquaculture est le secteur de production alimentaire qui connaît la croissance la plus rapide (d'environ 8 % par an depuis plusieurs décennies) et joue un rôle de plus en plus critique dans la sécurité alimentaire mondiale (FAO, 2020). La consommation de poissons d'élevage dépasse désormais celle de poissons issus de capture. L'aquaculture moderne implique donc des techniques de production massive, à l'image des élevages d'animaux terrestres. Ceci n'est pas sans conséquences, notamment en termes de développement de maladies. Les niveaux actuels de survie des poissons d'élevage, de l'éclosion jusqu'à la taille commerciale, sont de seulement 20 % en moyenne, et les experts estiment que les pertes mondiales de production dues aux maladies excèdent couramment 10 % (soit 6 milliards de dollars de pertes en 2014), ce qui en fait le défi majeur au développement de l'aquaculture (World Bank, 2014).

Le genre *Tenacibaculum* (famille des *Flavobacteriaceae*, phylum *Bacteroidetes*) comprend 32 espèces dont les noms ont été officiellement publiés (LPSN<sup>2</sup>), et dont au moins huit (*T. maritimum*, *T. discolor*, *T. soleae*, *T. dicentrarchi*, *T. gallaicum*, *T. finnmarkense*, *T. ovolyticum* et *T. piscium*) ont été signalées comme pathogènes des poissons marins, affectant une large gamme de poissons d'élevage à grande valeur ajoutée (e.g., saumon, bar, turbot, sole) avec une très large répartition géographique. Ces espèces sont responsables d'une maladie ulcéreuse (Figure 1) dénommée ténacibaculose (également connue sous le nom de flexibactériose) et associée à de graves épisodes infectieux conduisant à des pertes économiques importantes (Nowlan JP *et al.*, 2020). La plupart, sinon la totalité, des pays impliqués dans la production de poissons marins ont connu de graves flambées de ténacibaculose, récemment identifiée dans le projet MedAID H2020 comme la deuxième maladie la plus importante du bar dans les fermes piscicoles méditerranéennes (Muniesa *et al.*, 2020). À ce jour il n'existe pas de vaccins commerciaux, et les traitements reposent exclusivement sur des traitements antibiotiques ou l'utilisation de désinfectants chimiques.



Figure 1. Jeune bar (*Dicentrarchus labrax*) atteint de ténacibaculose et présentant une atteinte sur le flanc.

Dans la plupart des cas (y compris ceux de la littérature scientifique), l'isolement et l'identification bactérienne ne sont pas effectués et le diagnostic de ténacibaculose est uniquement basé sur les symptômes cliniques et des observations microscopiques des lésions. En effet, l'identification correcte des espèces impliquées dans la ténacibaculose est entravée par une combinaison de facteurs, en particulier la découverte récente d'une diversité insoupçonnée d'espèces pathogènes du genre *Tenacibaculum* qui n'a pas été suivie du développement des outils de diagnostic correspondants (Nowlan *et al.*, 2020).

Les méthodes modernes de typage bactérien reposent essentiellement sur des techniques moléculaires [e.g., séquençage de gènes multiples (MLST) ou du génome complet (WGS)] qui nécessitent du temps, une certaine technicité, ont un coût élevé, ne sont pas facilement automatisables et sont donc rarement applicables à du typage bactérien de routine. Le MALDI-TOF MS présentant bon nombre de ces qualités, il a donc été testé pour typer quelques bactéries, essentiellement pathogènes pour l'homme, comme *Klebsiella pneumoniae*, *Staphylococcus aureus*, *Pseudomonas aeruginosa* ou encore *Listeria monocytogenes*.

Outil innovant facile à utiliser, rapide, précis, bon marché et qui se prête bien à l'analyse à haut-débit, la spectrométrie de masse MALDI-TOF a ainsi révolutionné l'identification bactérienne dans les laboratoires de microbiologie médicale (Carbannelle & Nassif, 2011). Permettant d'identifier les micro-organismes au niveau de l'espèce par l'analyse des protéines totales (Croxatto *et al.*, 2012), cette technique

<sup>2</sup> LPSN - <https://lpsn.dsmz.de/search?word=tenacibaculum>, consulté le 20-01-2021.

est basée sur la génération de spectres de masse à partir de cellules entières (ou après extraction des protéines totales) et de leur comparaison à des listes de spectres de référence. Une fois les échantillons prêts, l'identification des espèces prend seulement quelques minutes. C'est ainsi que le MALDI-TOF MS a été intégré, au cours de la dernière décennie, dans les pipelines d'analyse de routine des laboratoires de microbiologie médicale pour l'identification microbienne, remplaçant les techniques biochimiques ou moléculaires traditionnelles. Cette technique trouve aussi de plus en plus sa place dans les laboratoires d'analyses vétérinaires.

Actuellement, le MALDI-TOF MS et les bases de données commerciales sont essentiellement focalisés sur l'identification taxonomique au niveau de l'espèce. Toutefois, la population d'individus constituant une espèce bactérienne est loin d'être homogène. Par conséquent, les espèces d'importance clinique sont subdivisées en groupes différents, par exemple, par leur degré de virulence, leur spectre d'hôte ou leur capacité de dissémination. Le but du typage bactérien est donc de discriminer des groupes de souches parmi d'autres et de comparer les isolats pour suivre la propagation de groupes spécifiques (e.g., clones épidémiques hautement virulents, cas sporadiques de moindre virulence) (Riley & Blanton, 2018). Le typage bactérien peut être utilisé pour réaliser des études épidémiologiques rétrospectives dans des laboratoires de recherche académique ou en routine dans des laboratoires d'analyses médicales, telle l'identification de clones épidémiques nécessitant la mise en œuvre rapide de mesures de contrôle. Par ailleurs, un des attraits du MALDI-TOF MS est que les spectres générés peuvent contenir beaucoup plus d'informations qu'il n'en faut pour l'identification des espèces bactériennes et qu'ils peuvent donc être utilisés à des fins de typage. Malheureusement, il ne semble pas possible d'établir un schéma universel de typage pour l'ensemble des espèces bactériennes, mais seulement des règles générales sur la manière d'établir et de valider un schéma (Spinali *et al.*, 2015). Il faut également souligner que la résolution du MALDI-TOF MS a certainement des limites et qu'il n'y a aucune garantie que les objectifs soient atteints (Kang *et al.*, 2017). Enfin, ceux qui cherchent à utiliser le MALDI-TOF MS pour effectuer du typage en routine ne peuvent probablement pas éviter d'analyser un grand

nombre de souches bien caractérisées et d'évaluer avec précision la valeur prédictive de leur schéma.

Bien que présentant de nombreux avantages, cette méthode nécessite des développements en vue d'une application en aquaculture. En effet, elle ne peut qu'identifier les espèces bactériennes pour lesquelles des spectres de référence préétablis sont présents dans les bases de données. Cette limite s'impose pour les nombreuses espèces pathogènes des poissons, pour la plupart absentes des collections de spectres fournies par les fabricants de spectromètres MALDI-TOF.

Cet article présente la stratégie mise en place pour identifier, par MALDI-TOF MS, les bactéries du genre *Tenacibaculum* et évaluer son potentiel pour le typage d'isolats appartenant à l'espèce *T. maritimum*. Ce retour d'expérience devrait permettre de guider le lecteur pour le développement d'une démarche similaire pour des espèces bactériennes intéressantes dans son domaine de recherche.

## MALDI-TOF MS : principe de fonctionnement

Il existe une très riche bibliographie sur le principe de fonctionnement du MALDI TOF-MS et cet article n'a pas vocation à le détailler (pour des revues, voir Singhal *et al.*, 2015 et Wieser *et al.*, 2012). Brièvement, un spectromètre MALDI TOF est composé de trois unités : (1) une source d'ions (MALDI), (2) un dispositif permettant de séparer les ions selon leur rapport masse / charge ( $m/z$ ) (TOF) et (3) un système de détection des ions. Dans le cas qui nous intéresse, l'échantillon analysé, ou analyte, se compose généralement de bactéries entières (d'extraits protéiques totaux) mélangées à une matrice favorisant l'ionisation, et les ions générés correspondent aux protéines bactériennes intactes ionisées. Les ions formés et détectés sont ceux chargés positivement (majoritairement de charge +1, c'est-à-dire ayant capté un proton). Ils sont accélérés sous l'action d'un champ électrique, puis passent dans un tube de vol. Dans ce tube, à charge égale (e.g.,  $z = +1$ ), ceux de petite masse sont les plus rapides, ceux de grosse masse sont les plus lents. Il existe une relation de proportionnalité entre la masse d'une protéine et son temps de vol (le TOF pour « time of flight »). C'est donc du temps de vol qu'est déduite la masse de chaque particule atteignant le détecteur. L'ensemble des ions détectés



va former un spectre caractéristique de l'échantillon analysé. Les appareils MALDI TOF-MS commerciaux dédiés à l'identification bactérienne sont capables de détecter des molécules de 20 à 20 000 m/z (ou Da pour tous ceux, majoritaires, de charge +1). Environ 50 % des protéines identifiées dans ces spectres MALDI-TOF sont des protéines ribosomiques, car elles sont majoritaires dans les cellules en croissance. La plupart d'entre elles ont un faible poids moléculaire, compatible avec la gamme de masse observable en MALDI-TOF, et elles possèdent également un p*K*i élevé, supérieur à 8, les rendant plus sensibles à l'ionisation (Ryzhov & Fenselau, 2001). Nous reviendrons ultérieurement sur ce point.

On représente classiquement le spectre d'un échantillon par un graphique avec, sur l'axe des abscisses, le rapport m/z et, sur l'axe des ordonnées, l'intensité du signal, c'est-à-dire le nombre total de chaque ion particulier détecté (Figure 2). Une fois le spectre d'un échantillon acquis, des logiciels permettent de comparer l'empreinte spectrale obtenue à toutes celles contenues dans une base de données.

La fiabilité et la précision du MALDI-TOF MS ont été vérifiées dans un certain nombre d'études ; les résultats d'identification obtenus reposent, pour une bonne partie, sur le nombre d'entrées de la banque de données utilisée (Singhal *et al.*, 2015 ; Hou *et al.*, 2019). À ce stade, les banques de données commerciales (ou bibliothèques de référence) contiennent essentiellement des spectres issus de micro-organismes d'intérêt clinique et peu de bactéries « exotiques », ce qui est le cas pour bon nombre d'espèces pathogènes des poissons.

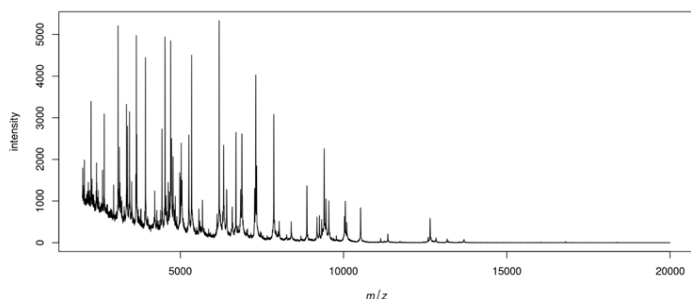


Figure 2. Exemple de spectre MALDI-TOF MS obtenu à partir de colonies entières de *Tenacibaculum maritimum*, suivi d'une étape d'extraction protéique par éthanol/acide formique. La matrice utilisée est l'acide -cyano-4-hydroxycinnamique.

## Création d'une bibliothèque de spectres de référence à façon

Il est heureusement possible de créer sa propre librairie de spectres de référence et de l'intégrer à une base de données commerciale (e.g., Bruker's MALDI Biotyper) ou à sa propre base de données. Comme il existe nombre de logiciels libres permettant de s'affranchir des outils commerciaux, nous avons choisi d'utiliser les outils MALDIquant, MALDIquant Foreign, et MALDIrppa (Gibb & Strimmer, 2012 ; Palarea-Albaladejo *et al.*, 2018) utilisant des scripts R. Comme évoqué précédemment, les résultats d'analyse MALDI-TOF MS d'un échantillon étant tributaires de la qualité de la banque de données (et dans une certaine mesure des logiciels d'analyse utilisés), la réalisation de cette (ou de ces) banque(s) est un prérequis nécessitant une attention particulière. Les souches types (celles à l'origine de la description de l'espèce) étant d'une importance fondamentale en systématique bactérienne, il nous est apparu indispensable d'intégrer celles de toutes les espèces du genre *Tenacibaculum* (LPSN). De plus, pour avoir une meilleure estimation de la variabilité des spectres obtenus, il nous semblait pertinent de générer plusieurs spectres de référence pour chaque espèce bactérienne particulièrement ciblée. Pour ce faire, il est nécessaire de sélectionner quelques souches représentatives de l'espèce bien caractérisées et de générer leurs spectres dans les mêmes conditions que celles utilisées pour les souches types. Pour réaliser ces spectres de référence, chaque souche type et chaque souche de référence est cultivée en triplicatas (trois répliques biologiques). Pour chaque réplique, 12 spectres sont enregistrés (4 dépôts indépendants sur la cible MALDI (répliques techniques) et 3 spectres acquis par dépôt). Il y a donc, en tout, 36 spectres pour chacune de ces souches. Après les contrôles qualité classiques, les spectres de bonne qualité retenus (24 spectres minimum) sont agrégés pour construire un spectre de référence moyen pour chaque souche.

## De l'intérêt de la génomique

Si l'utilisation du séquençage complet du génome (WGS) est encore difficilement envisageable pour effectuer du diagnostic de routine, en particulier sur un très grand nombre d'échantillons, il trouve, en

revanche, toute sa place pour caractériser finement les souches de référence. Cette stratégie permet de confirmer l'appartenance d'une souche à une espèce bactérienne (en utilisant par exemple des métriques basées sur l'identité nucléotidique moyenne ou ANI) et de préciser les positions taxonomiques des espèces les unes par rapport aux autres (e.g., phylogénies réalisées sur l'ensemble des gènes partagés entre souches ou génome central dans le taxon considéré). Ce travail permet également d'inclure ou d'exclure des souches pour générer les spectres de référence avec objectivité.

Dans le cadre du travail que nous avons réalisé sur les bactéries du genre *Tenacibaculum*, cette stratégie nous a permis i) d'infirmer l'existence de l'espèce *T. ascidiaceicola*, dont la souche type de cette espèce partage >98 % d'identité de séquence génomique avec l'espèce *T. discolor* (il aurait donc été compliqué de discriminer par MALDI-TOF MS une espèce qui n'en est pas une !); ii) de décrire la nouvelle espèce *T. piscium* (Olsen *et al.*, 2020); et iii) de confirmer l'identité taxonomique des souches intégrées dans notre stratégie pour réaliser la banque de spectres de référence.

L'obtention de génomes complets pour les espèces du genre *Tenacibaculum* nous a également permis de mettre en évidence l'implication de plusieurs espèces pathogènes des poissons étroitement apparentées (*T. dicentrarchi*, *T. finnmarkense* et *T. piscium*) dans des cas de tenacibaculose initialement attribués à *T. maritimum* (Bridel *et al.*, 2018; Olsen *et al.*, 2020). Ce résultat laissait, par ailleurs, entrevoir qu'il serait difficile de les discriminer par une approche classique, utilisant les spectres MALDI-TOF MS complets.

Enfin, comme évoqué précédemment, les spectres MALDI-TOF sont essentiellement constitués de protéines ribosomiques. Il est donc possible d'obtenir la séquence de ces protéines à partir des génomes afin d'y chercher des ressemblances et des différences. Cette approche visant à identifier des biomarqueurs permet de confirmer l'identification obtenue avec le spectre complet et de proposer un schéma de typage, du moins si du polymorphisme existe pour certains biomarqueurs.

## À la recherche de biomarqueurs

Si l'identification d'une espèce bactérienne par MALDI-TOF MS semble ne pas poser de grandes difficultés - sous réserve que la base de données interrogée soit complète et que les espèces considérées soient suffisamment distinctes les unes des autres -, il n'en est pas de même pour le typage bactérien. Un spectre MALDI-TOF MS contient généralement une centaine de pics, tout ou partie d'entre eux pouvant être utilisés pour du typage. Ce dernier peut donc soit exploiter le spectre entier, soit un sous-ensemble de pics ou encore un pic unique. Cependant, la capacité du MALDI-TOF MS à typer les souches dépend des caractéristiques intrinsèques de l'espèce bactérienne considérée. Il serait difficile, voire impossible, de capturer un polymorphisme suffisant pour des espèces très homogènes et présentant peu de diversité génétique, ou plus exactement protéique. Il n'est donc pas possible d'établir un schéma universel de typage, mais on peut s'appuyer sur des règles générales (Spinalli *et al.*, 2015; Sauget *et al.*, 2017) de manière à développer et à valider un schéma pour les espèces que l'on souhaite étudier.

Les données MALDI-TOF MS contiennent trois types d'informations : l'intensité du signal, la perte (et le gain) de signal et le *peak shift* (décalage du signal). Une variation d'intensité du signal peut être due à une modification d'expression de la protéine correspondante, liée aux conditions de culture, à l'âge de la colonie prélevée, etc. Une perte (ou un gain) de signal peut survenir en raison de l'absence (ou de la présence) d'un gène dans une souche donnée, mais également du mode de préparation de l'échantillon ou des conditions de culture. Ainsi, la variation d'intensité et l'apparition/disparition de pics donnent une information ambiguë sur le génotype d'un isolat. En revanche, un décalage de pic est la conséquence directe d'une modification de la séquence en acides aminés (polymorphisme) de la protéine correspondante. Ces modifications de composition en acides aminés d'une protéine se matérialisent par un changement de masse entraînant un décalage, en amont ou en aval, du pic théoriquement attendu pour la protéine correspondante. Avec la sensibilité des appareils MALDI-TOF MS, le changement d'un seul acide aminé est souvent suffisant pour observer dans les spectres un *peak shift*. Ils sont donc de

très bons candidats utilisables comme biomarqueurs pour effectuer un typage bactérien.

Sachant qu'environ 50 % des pics détectés dans un spectre MALDI-TOF MS correspondent à des protéines ribosomiques, il est relativement aisé d'identifier, à partir des génomes complets, des protéines ribosomiques présentant du polymorphisme susceptible d'entraîner un *peak shift*. On peut alors déduire les masses théoriques de ces isoformes protéiques et rechercher dans les spectres MALDI-TOF MS leurs positions relatives.

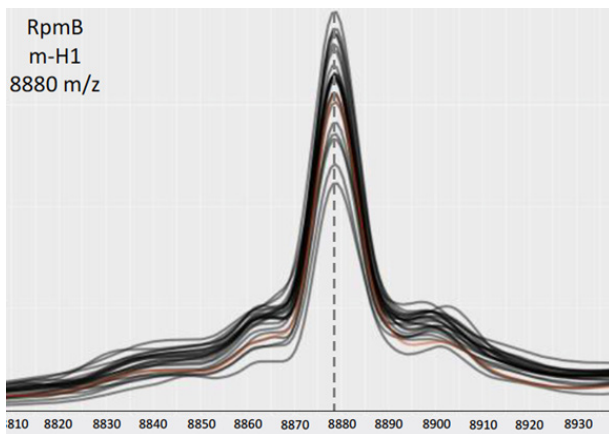
Nous avons effectué ce travail pour l'espèce *Tenacibaculum maritimum*. C'est l'espèce type du genre *Tenacibaculum*, sans doute celle la plus répandue dans les élevages et dont le plus grand nombre d'isolats était disponible. Nous avons donc sélectionné 24 souches dans la collection du laboratoire en les choisissant les plus diverses possibles (année d'isolement, origine géographique et poisson hôte). Nous avons séquencé les génomes de ces souches et réalisé en parallèle leurs spectres MALDI-TOF MS de référence. À partir des séquences obtenues, nous avons tout d'abord réalisé les vérifications d'appartenance à l'espèce (ANI, phylogénies utilisant le génome central...) et extrait les séquences déduites des protéines ribosomiques. Il était alors possible de les comparer et d'identifier celles ne présentant pas de polymorphisme (protéines ribosomiques monomorphes) et celles présentant des variations de séquence (protéines ribosomiques polymorphes). Les premières étaient de bonnes candidates biomarqueurs pour valider l'appartenance à l'espèce *T. maritimum* lors de l'identification des souches par MALDI-TOF MS. D'autre part, elles pouvaient servir de calibration interne pour l'ensemble du spectre obtenu. Les protéines ribosomiques polymorphes, quant à elles, devaient logiquement laisser apparaître un décalage de pic dans les spectres MALDI-TOF MS, et représentaient de bonnes candidates biomarqueurs pour développer une méthode de typage.

### Utiliser une combinaison de biomarqueurs pour le typage : développement d'une stratégie *Multi peak shift typing* (MPST)

Pour l'ensemble des protéines ribosomiques identifiées *in silico*, nous avons calculé les masses prédites (en prenant soin de conserver ou d'enlever la pre-

mière méthionine selon les résultats de prédiction obtenus par le logiciel « Terminator ») et recherché dans les spectres les pics correspondants. Nous avons retenu uniquement ceux présentant les critères stricts suivants : i) ils se situent dans la partie du spectre la plus discriminante (de 3 000 à 13 000 m/z) ; ii) ils correspondent aux masses des protéines ribosomiques prédites *in silico* avec une faible tolérance d'erreur (300 ppm) ; iii) ils sont présents dans tous les spectres MALDI-TOF des 24 isolats de référence ; iv) la plupart d'entre eux sont de haute intensité, même aux deux extrémités du spectre (bien que l'intensité soit globalement plus faible dans ces régions) ; v) ils ne sont pas perturbés par la présence d'autres pics très proches qui pourraient dégrader la fiabilité de détection des pics (en d'autres termes, les pics retenus devaient se trouver dans une fenêtre dépourvue de pics supplémentaires) ; et vi) pour les protéines polymorphes, nous avons privilégié les pics correspondant à des protéines monochargées (+1H). Ce travail nous a permis d'identifier 18 pics (soit 9 protéines différentes, certaines apparaissant en écho en fonction du nombre de charge +1H, +2H ou +3H) correspondant à des protéines ribosomiques monomorphes et 9 pics correspondant à des protéines ribosomiques polymorphes (Figure 3). Ces dernières présentaient de 2 à 5 isoformes différentes et ont servi à développer une stratégie de *Multi peak shift typing* (MPST). Nous avons, en effet, raisonné par analogie avec la stratégie de *Multi Locus Sequence Typing* (MLST) qui repose sur une combinaison de marqueurs correspondant à des séquences nucléiques, généralement des 7 gènes suffisamment conservés dans une espèce, mais présentant un certain polymorphisme permettant de discriminer les souches entre elles. Nous avons donc codé de manière arbitraire (1, 2, 3...) les différentes isoformes des protéines polymorphes retenues (à la manière des *allele type* ou AT dans l'approche MLST) et les avons agrégées : le résultat ou MALDI-type correspond à la combinaison unique d'isoformes pour chacun des 9 biomarqueurs présents dans une souche donnée (à la manière des *sequence type* ou ST correspondant à une combinaison unique d'ATs dans l'approche MLST). Cette stratégie originale permettait, par ailleurs, d'utiliser, dans le cadre des données MALDI-TOF MS, les outils informatiques (eBURST, Split tree) développés à l'origine pour re-

A)



B)

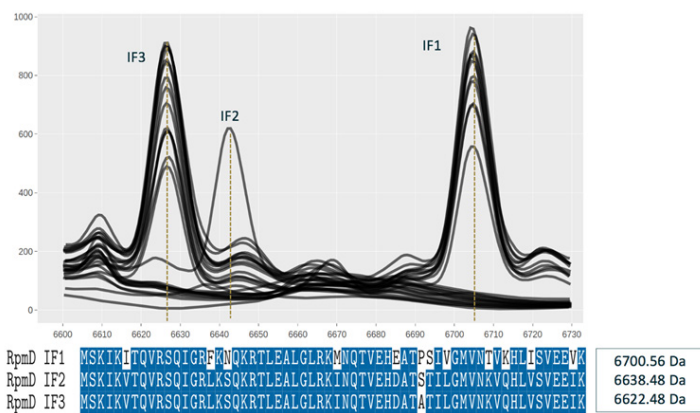


Figure 3. Exemple de pics correspondant à des protéines ribosomiques de *Tenacibaculum maritimum* observées dans les spectres MALDI-TOF MS. A) protéine RpmB, monomorphe monochargée (masse prédite 8880 Da après retranchement de la méthionine en N-terminal). B) Alignement des trois isoformes (IF1, IF2 et IF3) de la protéine polymorphe RpmD à partir des séquences extraites de 24 génomes représentatifs de l'espèce *T. maritimum* (les masses prédites sont indiquées) et pics correspondant à ces trois isoformes, identifiés dans les spectres MALDI-TOF MS obtenus à partir des mêmes souches.

présenter les données MLST. Ces outils ont permis de visualiser les liens, dont ceux issus de la recombinaison, entre souches et de les regrouper en complexe clonaux (CC).

À partir des données MALDI-TOF MS obtenues sur 131 isolats de terrain appartenant à l'espèce *T. maritimum*, notre approche MPST nous a donc permis de discriminer 2 à 5 isoformes/pic (équivalent aux ATs), 20 combinaisons différentes, ou MALDI-types (équivalent aux STs), et 4 groupes (équivalent aux CCs) ainsi qu'un singleton (Bridel *et al.*, 2020).

## De l'intérêt d'un site web intégrant une base de données et des outils d'analyse

Un des avantages majeurs de la MLST par rapport à toutes les approches traditionnelles a été le développement de bases de données regroupant pléthore d'espèces bactériennes, leurs différentes formes alléliques et des outils d'analyse. Ces bases de données, interrogeables à distance (cf. <https://pubmlst.org/>), permettent d'effectuer des suivis épidémiologiques en temps réel et à l'échelle mondiale ainsi que la comparaison fine de souches. Elles ont largement contribué au succès de cette approche.

Nous avons donc développé un outil en ligne, MALDIquant-TypeR, qui intègre les spectres de référence ainsi que les algorithmes élaborés durant ce projet (<http://genome.jouy.inra.fr/shiny/malდიquanttyper/>). Il permet de charger les spectres d'un isolat, puis de les comparer à l'intégralité de ceux des espèces du genre *Tenacibaculum* présentes dans notre base de données. Il intègre des outils d'identification basés sur la comparaison du spectre entier d'un isolat de terrain, mais également basés sur des biomarqueurs spécifiques identifiés pour différentes espèces pathogènes du genre. Si l'échantillon est identifié comme appartenant à l'espèce *T. maritimum*, l'outil applique le schéma de typage décrit dans cet article et propose une attribution pour chacun des 9 biomarqueurs retenus. Lorsqu'ils sont tous identifiés dans un échantillon, l'outil affiche le MALDI-type correspondant.

## Conclusion

Les outils de diagnostic sont essentiels pour une détection précoce des maladies, en particulier en l'absence de signes cliniques spécifiques et parce que l'agent causal pourrait être allochtone, émergent ou la maladie d'origine multifactorielle (de Lorgeril *et al.*, 2018). Ainsi, des approches non ciblées (à la différence de la plupart des techniques employant des réactions de PCR) sont à privilégier, notamment pour les pathologies en aquaculture, aussi bien des mollusques que des poissons (Moussa *et al.*, 2021). La caractérisation d'isolats par MALDI-TOF MS est une approche générique, non ciblée et très efficace lorsqu'il s'agit d'étudier un grand nombre d'échantillons. La rapidité d'acquisition des données et de leur ana-



lyse ainsi que la fiabilité des informations qu'elles contiennent sont des avantages indéniables en comparaison avec d'autres méthodes plus traditionnelles, comme les tests biochimiques ou la sérologie, ou plus modernes et ciblées comme la MLST. Comparée à cette dernière, l'approche par MALDI-TOF sacrifie un peu en précision pour gagner grandement en rapidité et en coût. C'est donc l'outil idéal pour réaliser des primo-analyses sur un grand nombre d'isolats afin de sélectionner ceux susceptibles d'être étudiés plus en détail (par exemple, par séquençage de génomes complets).

Il faut cependant garder à l'esprit que les spectres générés par MALDI-TOF MS sont des données analogiques complexes par nature (soumises au bruit de fond intrinsèque des appareils, à la méthode de culture des bactéries, à la préparation des échantillons, à la matrice d'ionisation utilisée, etc.). Elles se différencient donc des données de séquençage qui sont de nature numérique (présence de la base A, T, G ou C à la position n de la séquence) et nécessitent un traitement du signal adapté. Par ailleurs, l'emploi en routine du MALDI-TOF MS s'est, jusqu'à présent, essentiellement focalisé sur des espèces bactériennes bien caractérisées et d'importance clinique. Les banques de données commerciales ne contiennent pas ou peu d'espèces bactériennes « exotiques » ; il est donc nécessaire de construire sa propre bibliothèque de spectres de référence quand on s'intéresse à des taxons bactériens insolites, non encore inclus dans les banques de données existantes. Si l'identification de l'espèce ne semble pas poser de problèmes majeurs (sous réserve des éléments évoqués précédemment), l'emploi du MALDI-TOF MS pour entreprendre du typage est une entreprise plus ardue. Elle nécessite de combiner des approches de génomique analytique pour identifier des biomarqueurs susceptibles d'être observables dans les spectres. Ces étapes sont chronophages,

et l'implémentation d'algorithmes de traitements dédiés nécessite de réelles compétences en bioinformatique. Cependant, passées ces étapes de développement, la stratégie se révèle redoutablement performante. Elle permet de cribler à grande vitesse et à très faible coût des centaines ou milliers d'isolats bactériens. Un autre avantage non négligeable de la méthode est qu'elle fonctionne aussi bien sur des cultures fraîches que sur des cultures conservées sous forme inactivée jusqu'à 15 jours dans une solution d'éthanol, ce qui permet leur acheminement vers des laboratoires équipés de spectromètres de masse MALDI-TOF éloignés des sites de production. Cette facilité de transport favorise grandement les envois internationaux, les bactéries fixées en éthanol n'étant pas soumises à toutes les formalités administratives liées au transport de matériel biologique vivant, en particulier quand il s'agit d'agents pathogènes. Depuis la publication de ce travail, en 2020 (Bridel *et al.*, 2020), nous avons pu caractériser, par cette technique, plusieurs centaines d'isolats appartenant au genre *Tenacibaculum*, en provenance de Norvège, du Chili et de Polynésie française. Enfin, le déploiement de l'outil en ligne, MALDIquant-TypeR, devrait faciliter, dans un proche avenir, des études épidémiologiques à large échelle.

On peut enfin noter que des approches purement bioinformatiques (e.g., analyses du panspectrome, clustering et recherche de co-occurrences multiples de pics intégrant des méthodes statistiques) pourraient peut-être éviter, dans le futur, de passer par la recherche systématique de biomarqueurs pour réaliser du typage, ou du moins permettre de discriminer des groupes à l'échelle infra-espèce (Giraud-Gatineau *et al.*, 2020). Cependant, ces approches nécessitent de très grands jeux de données (au minimum 600 spectres de référence par espèce ciblée) et ne seront sans doute pas applicables à l'ensemble des espèces bactériennes. ■

# Remerciements

Je tiens à remercier particulièrement Jean-François Bernardet pour avoir construit et entretenu, pendant quarante ans, la souchothèque de l'équipe Infection et Immunité des poissons de l'Unité Virologie et Immunologie Moléculaires, pour ses précieux conseils et pour sa contribution à l'amélioration de ce document ; Tatiana Rochat pour sa contribution à l'amélioration de ce document ; Frédéric Bourgeon et l'entreprise Bio Chêne Vert pour avoir réalisé l'ensemble des spectres MALDI-TOF analysés dans le travail cité ; ce dernier a fait l'objet de la thèse CIFRE de Sébastien Bridel en partenariat avec l'entreprise Labofarm (groupe Finalab). Ce travail a reçu le soutien financier Carnot France Futur Élevage F2E 2017, projet « FishPathoMaldi ».

# Références

Bridel S., Olsen A.B., Nilsen H., Bernardet J.F., Achaz G., Avendaño-Herrera R. And Duchaud E., 2018. Comparative genomics of *Tenacibaculum dicentrarchi* and "*Tenacibaculum finnmarkense*" highlights intricate evolution of fish-pathogenic species. *Genome Biol Evol.* 2018 Feb 1;10(2):452-457. doi: 10.1093/gbe/evy020. PMID: 29360975; PMCID: PMC5793721.

Bridel S., Bourgeon F., Marie A., Saulnier D., Pasek S., Nicolas P., Bernardet J.F. and Duchaud E., 2020. Genetic diversity and population structure of *Tenacibaculum maritimum*, a serious bacterial pathogen of marine fish: from genome comparisons to high throughput MALDI-TOF typing. *Vet Res.* 2020 May 7;51(1):60. doi: 10.1186/s13567-020-00782-0. PMID: 32381115; PMCID: PMC7204230.

Carbonnelle E. And Nassif X., 2011. Utilisation en routine du MALDI-TOF-MS pour l'identification des pathogènes en microbiologie médicale [Applications of MALDI-TOF-MS in clinical microbiology laboratory]. *Med Sci (Paris).* 2011 Oct;27(10):882-8. French. doi: 10.1051/medsci/20112710017. Epub 2011 Oct 21. PMID: 22027426.

Croxatto A., Prod'homme G. And Greub G., 2012. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol Rev.* 2012 Mar;36(2):380-407. doi: 10.1111/j.1574-6976.2011.00298.x. Epub 2011 Aug 22. PMID: 22092265.

De Lorgeril J., Lucasson A., Petton B., Toulza E., Montagnani C., Clerissi C., Vidal-Dupiol J., Chaparro C., Galinier R., Escoubas J.M., Haffner P., Dégremont L., Charrière G.M., Lafont M., Delort A., Vergnes A., Chiarello M., Faury N., Rubio T., Leroy M.A., Pérignon A., Régler D., Morga B., Alunno-Bruscia M., Boudry P., Le Roux F., Destoumieux-Garzn D., Gueguen Y. and Mitta G., 2018. Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters. *Nat Commun.* 2018 Oct 11;9(1):4215. doi: 10.1038/s41467-018-06659-3. PMID: 30310074; PMCID: PMC6182001.

FAO, 2020. La situation mondiale des pêches et de l'aquaculture 2020 : La durabilité en action. <http://www.fao.org/3/ca-9231fr/CA9231FR.pdf>.

Gibb S. and Strimmer K. 2012. MALDIquant: a versatile R package for the analysis of mass spectrometry data. *Bioinformatics.* 2012 Sep 1;28(17):2270-1. doi: 10.1093/bioinformatics/bts447. Epub 2012 Jul 12. PMID: 22796955.

Giraud-Gatineau A., Texier G., Garnotel E., Raoult D. and Chaudet H., 2020. Insights into subspecies discrimination potentiality from bacteria MALDI-TOF Mass Spectra by using data mining and diversity studies. *Front Microbiol.* 2020 Aug 13;11:1931. doi: 10.3389/fmicb.2020.01931. PMID: 32903575; PMCID: PMC7438549.

Hou T.Y., Chiang-Ni C. and Teng S.H., 2019. Current status of MALDI-TOF mass spectrometry in clinical microbiology. *J Food Drug Anal.* 2019;27:404-14.

Kang L., Li N., Li P., Zhou Y., Gao S., Gao H. et al, 2017. MALDI-TOF mass spectrometry provides high accuracy in identification of *Salmonella* at species level but is limited to type or subtype *Salmonella* serovars. *Eur J Mass Spectrom (Chichester).* 2017;23:70-82.

Moussa M., Cauvin E., Le Piouffle A., Lucas O., Bidault A., Paillard C., Benoit F., Thuillier B., Treilles M., Travers M.A. and Garcia C., 2021. A MALDI-TOF MS database for fast identification of *Vibrio* spp. potentially pathogenic to marine mollusks. *Appl Microbiol Biotechnol.* 2021 Feb 15. doi: 10.1007/s00253-021-11141-0. Epub ahead of print. PMID: 33590268.

Muniesa A., Basurco B., Aguilera C., Furones D., Reverté C., Sanjuan-Vilaplana A., Jansen M.D., Brun E. And Tavorpanich S., 2020. Mapping the knowledge of the main diseases affecting sea bass and sea bream in Mediterranean. *Transbound Emerg Dis.* 2020 May;67(3):1089-1100. doi: 10.1111/tbed.13482. Epub 2020 Jan 30. PMID: 31960605.

Nowlan J.P., Lumsden J.S. and Russell S., 2020. Advancements in characterizing *Tenacibaculum* infections in Canada. *Pathogens.* 2020 Dec 8;9(12):1029. doi: 10.3390/pathogens9121029. PMID: 33302445; PMCID: PMC7763822.

Olsen A.B., Spilsberg B., Nilsen H.K., Lagesen K., Gulla S., Avendaño-Herrera R., Irgang R., Duchaud E. and Colquhoun D.J., 2020. *Tenacibaculum piscium* sp. nov., isolated from skin ulcers of sea-farmed fish, and description of *Tenacibaculum finnmarkense* sp. nov. with subdivision into genomovars *finnmarkense* and *ulcerans*. Int J Syst Evol Microbiol. 2020 Dec;70(12):6079-6090. doi: 10.1099/ijsem.0.004501. PMID: 33079030.

Palarea-Albaladejo J., Mclean K., Wright F., Smith D.G.E., 2018. MALDIrppa: quality control and robust analysis for mass spectrometry data. Bioinformatics. 2018;34:522-3.

Riley L. and Blanton R., 2018. Advances in molecular epidemiology of infectious diseases: definitions, approaches, and scope of the field\*. microbiol-spec 6(6): doi:10.1128/microbiol-spec.AME-0001-2018.

Ryzhov V. and Fenselau C., 2001. Characterization of the protein subset desorbed by MALDI from whole bacterial cells. Anal Chem 73:746-750.

Sauget M., Valot B., Bertrand X. and Hocquet D. Can MALDI-TOF Mass Spectrometry reasonably type bacteria? Trends Microbiol. 2017;25:447-55.

Singhal N., Kumar M., Kanaujia P.K., Viridi J.S., 2015. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front Microbiol. 2015;6. doi:10.3389/fmicb.2015.00791.

Spinali S., Van Belkum A., Goering R.V., Girard V., Welker M., Van Nuenen M. et al., 2015. Microbial typing by matrix-assisted laser desorption ionization-time of flight mass spectrometry: do we need guidance for data interpretation? J Clin Microbiol. 2015;53:760-5.

Wieser A., Schneider L., Jung J. and Schubert S., 2012. MALDI-TOF MS in microbiological diagnostics-identification of microorganisms and beyond (mini review). Appl Microbiol Biotechnol. 2012 Feb;93(3):965-74. doi: 10.1007/s00253-011-3783-4. Epub 2011 Dec 25. PMID: 22198716.

World Bank, 2014. Reducing disease risks in aquaculture. World Bank Report #88257-GLB.



Cet article est publié sous la licence Creative Commons (CC BY-SA). <https://creativecommons.org/licenses/by-sa/4.0/>

Pour la citation et la reproduction de cet article, mentionner obligatoirement le titre de l'article, le nom de tous les auteurs, la mention de sa publication dans la revue « NOV'AE », la date de sa publication et son URL.