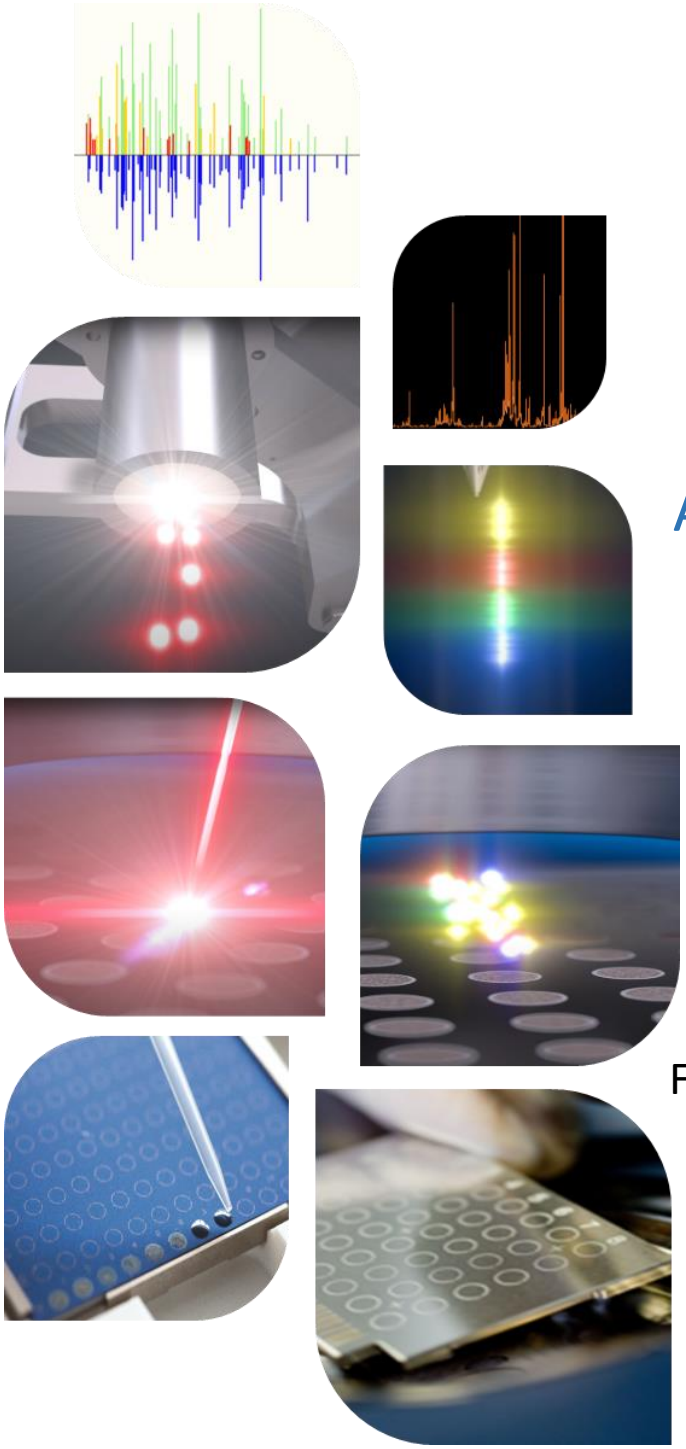

STANDARDS & GUIDELINES:

GENERATION AND ASSESSMENT OF MALDI-TOF MS FINGERPRINT SPECTRA AND MANAGEMENT OF FINGEPRINT LIBRARIES

FOR APPLICATION TO AUSTRALIAN BIOSECURITY DIAGNOSTIC LABRATORIES

Version 1 May 2024

Authors

Christopher Nellessen
Mass Spectrometrist
Science and Surveillance Group |Biosecurity Plant Division
Australian Government Department of Agriculture, Fisheries and Forestry

David Nehl
Director - Plant Pathology and Botany Section
Science and Surveillance Group |Biosecurity Plant Division
Australian Government Department of Agriculture, Fisheries and Forestry

Toni Chapman
Senior Research Scientist
Plant Biosecurity Research and Diagnostics
NSW Department of Primary Industries

Dominie Wright
Principal Laboratory Scientist
DPIRD Diagnostic Laboratory Services
Department of Primary Industries and Regional Development.

Contents

INTRODUCTION

The Department of Agriculture, Fisheries and Forestry (DAFF) invested in a MALDI-ToF MS pilot project to investigate the challenges and usefulness of the technology's inclusion would bring to the Science and Surveillance Group's (SSG) routine plant pathology biosecurity diagnostics. After a series of informative workshops regarding the use and issues of the MALDI-ToF MS pilot project based upon the resulting feedback, the Plant Health Committee recommended establishment of a National MALDI-ToF MS Steering Committee to oversee a national approach to the development and sharing of MALDI-ToF MS diagnostic reference libraries. Chaired by DAFF's Kath De Boer, the committee includes other DAFF members and representatives from other biosecurity departments using MALDI-ToF MS, namely, Western Australia Department of Primary Industries and Regional Development as well as New South Wales Department of Primary Industries. Observers representing other biosecurity departments are also included in the committee's membership. The technical expertise of the included members ensures the established standards are of sufficient quality for widespread application.

Mass spectrometry (MS) facilitates productive scrutiny of an unknown target's identity and nature by exactingly measuring the mass of atoms and molecules making up that target. This technique has been leveraged in many disciplines to great success, with notable achievements in biomolecule analysis, making it an essential tool in proteomic, metabolomic, and lipidomic studies while paving the way for its introduction into modern biosecurity diagnostics. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-ToF MS) represents one of the best mass spectrometry instruments to adopt for biosecurity for the same reasons that it has been widely implemented for other diagnostic demands—it is easy to use, rapid, and readily commercially available and supported.

Mass spectrometry analysis represents another molecular approach fundamentally different from genetic sequencing techniques that maintains the reproducibility, traceability, and reliability which is important in any modern analytic technique. MALDI-ToF MS diagnostics involves comparing unknown sample results to reference mass spectra (fingerprints) from verified samples to generate identifications. As MALDI-ToF targets can be prepared quickly and easily from cultures, or even directly from infected material, the technique has significant potential to improve biosecurity diagnostics. Bacteria are the most readily identified target using MALDI-ToF MS thanks to the existence of extensive, accessible bacteria fingerprint libraries—though notable taxonomic gaps remain as development efforts concentrated on human and veterinary pathology. Fungi also feature similarly available fingerprint libraires, though they are less extensive, limiting the breadth of genera/species they can identify. Application of MALDI-ToF MS for other targets is restricted by the lack of available fingerprint libraries, though the investigative power of the technique remains the same.

MALDI-ToF MS represents a significant modernisation of the traditional plant / biosecurity diagnostic laboratory, in the same vein as genetic sequencing, but with a much lower cost

for individual sample application and a much higher throughput than Sanger or nanopore sequencing techniques. The technique can be applied early in the plant diagnostic workflow and is easily used on multiple targets—leading to faster diagnostic results, while enabling better targeting of sequencing techniques for follow-up testing.

Effective adoption of the technology is generally restricted by the breadth of the pathogen libraries available to the diagnosticians using the technique. Manufacturer recommendations regarding fingerprint creation are limited, and their quantitative quality assessment focuses only on single aspect of the fingerprint data, while leaning heavily on qualitative assessment to ensure usefulness. Subsequently, production of supplementary fingerprint libraries is difficult for most laboratories. A consolidated approach paired with constructive assessment criteria would enable superior standardisation while having wider implications regarding MALDI-ToF MS usage.

Standardisation of reference library development is necessary to ensure quality and downstream accuracy and reliability of the technology for new or emerging pathogens as well as for additional targets not supported by currently available libraries. The introduction of additional quantitative elements within the quality assessment of reference spectra would improve transparency, reproducibility, and consistency of quality expectations, while enabling superior quality assurance during the development process and post-production library quality review. The new standards would consist of a collection of quantitative spectral data analysis procedures paired with a simple qualitative assessment to generate an overall quality judgement that enables effective reference spectra screening during the development of reference spectra libraries. Also included are a set of guidelines to enable acceptable result production without significant impediment to investigative or diagnostic analysis.

This document is directed at the adoption of quality reference library development to further the scope of MALDI-ToF MS diagnostics for general diagnostics and biosecurity across all Australian jurisdictions. The assessment criteria described in this document supports laboratory accreditation when combined with standard quality assurance practices.

How to use this document

This document is not a methods manual. Laboratories must ensure they are familiar with appropriate methodology to conduct MALDI-ToF MS testing. This document focuses exclusively on the Bruker MALDI Biotyper® and will not apply directly to bioMérieux/Shimadzu or other MALDI-ToF MS instruments due to the different way that those commercial systems construct fingerprints. Users of alternate MALDI-ToF MS systems must consider how to develop their own equivalent quality standards for library development. It is important to note that there is no direct conversion process allowing such libraries to be accessible between commercial MALDI-ToF MS systems. Full documentation of alternate system fingerprint library standards requires full documentation for quality control points within their respective MALDI-ToF MS system.

This document does not replace the quality system of the laboratory it is used in and should be used in conjunction with and addition to the quality assurance and quality control requirements of that system.

This document provides minimum standards that must be used to ensure commonality reference libraries being developed across jurisdictions. The document also includes a set of “guidelines” that exceed the standards and are provided as a recommendation.

This document is intended for ensuring confidence and commonality in diagnostic applications. It should be considered optional whether laboratories choose to apply these standards to their research and development activities, though lack of use may impact fingerprint sharing. This document is not intended to limit the use of MALDI-ToF MS fingerprint library development in emergency scenarios such as disease/pest outbreaks where bespoke fingerprint development may be necessary to enable rapid response if the standard is deemed unsuitable. Such use is considered research in the context of this document.

Glossary & Abbreviations used in this document

Term	Definition
Biotyper®	The brand name of the Bruker MALDI Biotyper® Sirius RUO instrument system that combines a MALDI-ToF MS with software featuring a spectrum matching algorithm
BTS	Bruker bacterial test standard, a commercially available, traceable, quality control and calibration material made up of a carefully manufactured, modified *E. coli* extract with a known, characteristic mass spectrum.
Bruker	Bruker Pty Ltd. The MALDI-ToF MS vendor and the source of the commercial bacterial and fungi libraries
Guideline	A recommended beneficial activity that exceeds the standard
Intensity	The value represented by arbitrary units (a.u.) on the y-axis of a mass spectrum. Not a quantitative measure of the amount of a particular parent compound producing an ion in the sample, rather a measure of the abundance of detected ions for a particular m/z *relative* to the rest of the detected ions in the mass spectrum.
Library	A collection of reference spectra that can be compared to the spectrum of a tested sample to generate close matches and inform the identification of that sample. The identification power of a library is directly related to the number of reference spectra it includes.
MALDI	Matrix assisted laser/desorption ionization, the part of the instrument's analysis that prepares a sample for detection by ionizing the mixture of the sample and matrix using a high intensity pulse UV laser
Matrix	The light-absorbing chemical that allows for ionization and therefore detection of the target sample. It is allowed to mix with or is layered on top of the sample to ensure this occurs during the MALDI process
Metadata	Data that describe other data, structured reference data that helps to sort and identify attributes of the information it describes. Metadata summarises basic information about data, which can make it easier to find, classify, use and reuse particular instances of data.
MS	Mass spectrometer, the overall instrument as well as the analytical method that returns results indicating the exact masses of the different molecules that makeup of a sample. The MALDI-ToF MS gives graphically represented results where all the molecules in the sample are plotted by their mass and amount detected, allowing for similar results to be compared.
m/z	Mass-to-charge ratio, the measured value for a mass spectrometer. This information is represented graphically by the position of a peak on the x-axis of a mass spectrum.
Noise	Noise is undesirable background signal in a mass spectrum that can arise chemically or instrumentally. This is distinct from contamination, which can also produce undesirable peaks within a mass spectrum. Chemical noise is random signal caused by molecular interactions and can arise due to analyte or detector interaction with the system environment. Instrument noise is caused by instrument itself and is of low impact due to modern analytical capacity. Random noise is most easily reduced in a mass spectrum through ensemble averaging, like that performed automatically during Peak MALDI data collection.
Quality assurance	The infrastructure in place to ensure quality through application of best practices, quality controls, and process management. This provides confidence that results are of acceptable quality through prevention of quality issues ensuring both internal management and external customers, government agencies, regulators, certifiers, and third parties expectations are fulfilled.
Reference Spectrum or MSP (Main Spectra Profile) or Fingerprint	A reference spectrum, MSP, or fingerprint, is the mass spectrum information, including peak position and height, extracted from multiple spectra produced from a known sample and compiled into a single set of data. Reference spectrum is the terminology used most widely, while MSP is the term used by Bruker for these spectra, and fingerprint is commonly used for layman for simple understanding of its use. The mass

Term	Definition
	spectrum information mostly corresponds to proteins and peptides and remains consistent for specimens of the same species over set mass ranges.
Sample	In the context of this document, sample means any biological material, including but not limited to, bacteria and other microorganisms, plant, animal, environmental, or cell lines containing copies of the original sample components. A sample may refer also to derivatives from these materials.
S/N	Signal-to-noise ratio, an analytical measure of quality that for MALDI-ToF MS is based upon the ability to distinguish meaningful signal of an analyte ion from spectrum noise.
Spectrum	Spectrum (*pl.* spectra) is the graphical output of a mass spectrometer that features peaks which correspond to the exact masses divided by charge of detected molecules on the x-axis and heights that correspond to amount of that molecule which was ionized on the y-axis
Standard	A repeatable, harmonised, agreed, and documented way of doing something. Standards contain technical specifications or other precise criteria designed to be used consistently. Defined by Standards Australia as documents that set out specifications, procedures and guidelines that aim to ensure products, services, and systems are safe, consistent, and reliable
Target Plate	The sample holder used during analysis that features outlined targets that are filled with sample and matrix. Disposable target plates use a proprietary base material to aid ionization and detection, as well as hydrophobic target outlines to concentrate the sample during drying, making them the preferred consumable for use during analysis. Reusable target plates require cleaning with trifluoracetic acid and offer poorer results over time but has the advantage of lower costs after initial capital expenditure.
ToF	Time-of-Flight, the part of the instrument's analysis that separates the different molecules in the sample by their exact mass by allowing them to drift through a vacuum tube and also includes the detector
Validation	Determination and confirmation of the performance characteristics of a process or test through rigorous testing
Verification	Confirmation of the characteristics of a pre-validated process in a particular environment, such as a laboratory or for a particular sample or instrument

MALDI-ToF Mass Spectrometer Maintenance and Health

Standardisation of reference spectra or fingerprints requires initial instrument parity in the context of its operational health and expected ability to generate quality results. While regular quality control measures assist in this, there are several more general instrument checks that allow for instrument health traceability and aid troubleshooting diagnostics. As these aspects play a role during reference spectrum development, both standards and guidelines are outlined for application to the instrument and quality assurance management practices.

QC/Calibration records

Operation of MALDI-ToF MS in a traceable, quality assured manner requires regular calibration and quality control procedures. Generation of fingerprints requires the use of calibration to ensure mass accuracy, with a maximum acceptable mass shift of 300 ppm for each calibration peak, but calibration assessment provides no information regarding the health of the instrument over time. Rather, the end users should implement a recording system to ensure that the calibration and quality control measures fit within known, expected boundaries and that potential calibration or QC issues can be detected when they occur.

The key quality control measure that should be recorded is derived from measuring the BTS during automated analysis. We recognize during the automated QC process that multiple quality control measurements are performed and that these are not directly visible to the end user without exploration of the data files through CompassExplorer. Further, the analysis of the BTS spot that generates a matching score occurs after the automated QC analysis and has no impact on this automated QC check. Given the difficulty of accessing and interpreting the latter data, monitoring the final matching score of the BTS spot is recommended as an alternative. Manual analysis of BTS lacks the inherent quality control guarantees of the automated QC method, but can still generate a matching score to *E. coli* that can be used to monitor instrument health. Alterations to the default automated data collection method or BTS validation method are restricted to changes that do not alter the number of distinct data collections or tests performed during the quality control analysis. Changes to alter data collection quality through expanding the data collection time or mass information gathered by increasing or decreasing the number of laser or raster shots is not an alteration to the number of collections, and therefore is an acceptable change, though such method changes should be performed on the validation method only to bring it more in line with an altered data collection method.

Standard – BTS quality control check records

- **Traceability of proper mass accuracy of the instrument requires records of quality control results. This may be the matching *E. coli* score for the quality control material following an automated data collection or the analysis of matches for each individual quality check performed during the automated collection through CompassExplorer file navigation.**

- **Records must contain analysis of mean and standard deviation and record how individual quality control results compare to these values.**

Guidelines

- A Levey-Jennings[1] or equivalent graph is recommended to assist in visualizing trends in the quality control and catch possible instrument measurement drifts that may require addressing.
- Use of the automated quality control procedure for BTS analysis is recommended to ensure instrument results are of high-quality.
- Use of Peak MALDI is recommended for more consistent results—Peak MALDI is recommended to be used with a validation method featuring additional shots and/or laser rastering to account for the more extensive matrix depletion caused by Peak MALDI.

Calibration records are important to monitor the ongoing health of the instrument. The calibration can be performed automatically or manually and doing so is required before collecting fingerprint data to ensure high mass accuracy is achieved. Manual calibration requires manual assessment of the results produced by the BTS to ensure mass shifts are within the allowable 300 ppm, as recommended by the manufacturer, and results should be recorded. All calibration records, regardless of collection method, should be organized and labelled by date and are recommended to include the three variables that define the mass adjustments performed by the instrument during the data collection to fit the known calibrant's peaks to the current detection results. Following a successful calibration, these three calibration constants, C_0, C_1, and C_2 can be found by navigating to the calibration tab on the FlexControl software and clicking the properties button. The three constants will be listed and can be recorded in a separate spreadsheet for record purposes. The average value and standard deviation for each parameter should be monitored, and each individual value should be recorded, graphed, and compared to other results through a moving average trend line to monitor variation over time.

Standard – calibration records

- **Traceability of proper mass accuracy of the instrument requires records of each of the three calibration constants.**
- **Each record must contain an analysis of the mean and standard deviation.**
- **Individual calibration constant results should be compared to recorded mean and standard deviation values.**

Guidelines

- A Levey-Jennings or equivalent graph is recommended to assist in visualizing trends in the three different calibration constants and catch potential instrument mass accuracy drifts that may require addressing.

Servicing/maintenance

Result quality is heavily impacted by MALDI-ToF instrument health, with a broad and noticeable improvement in nearly all measured variables correlating directly with instrument cleanliness and proper instrument maintenance. Therefore, instruments must be serviced regularly to address ongoing issues, detect potential problems, and ensure the highest quality fingerprint data can be generated. Servicing and preventative maintenance by support staff, such as Bruker® engineers, should be based upon instrument usage rates, but it is routinely required to ensure that fingerprint data quality remains consistent.

Standard – servicing

- **At minimum, servicing and preventative maintenance by trained support staff, such as Bruker® engineers, is required once per year and should be recorded in conjunction with calibration records**

Guidelines

- The rate at which manufacturer support staff are used to service the instrument should be based upon the instrument's routine workload and use. This may necessitate either additional visits by the support staff each year, or careful internal cleaning by trained staff to maintain the quality needed to enable high-quality reference spectra data collection.

Source Cleaning

As result quality is directly affected by instrument cleanliness, automated source cleaning should be performed regularly to ensure that high-quality fingerprint data can be collected. The Bruker® line of MALDI-ToF MS instruments estimates source contamination, and this should guide the use of the automatic source cleaning feature. Failure to keep the source clean may result in BTS quality control or calibration failures, in addition to potentially reducing the diagnostic information obtained from other analysed samples.

Standard – source cleaning

- **The source should be at ≤50% contamination by the instrument estimate before beginning fingerprint data collection.**

Guidelines

- The automated source-cleaning procedure should be performed before every fingerprint data collection.
- If collection of fingerprint data is extensive (e.g., more than 50 spots) or the samples have a higher risk of contaminating the source (e.g., older invertebrate samples), the source contamination level should be checked before proceeding with each run and source-cleaning performed if required.

Detector Checks

The signal and quality of the result signal are both related to the sensitivity of the detector in the MALDI-ToF MS instrument being used. Monitoring the health of the detector over time is essential for early discovery of sensitivity issues that may affect fingerprint generation. Regular detector checks enable laboratories to find and address sensitivity issues by setting the detector voltage to the appropriate level.

Standard – signal detection

- **A detector check should be performed at least quarterly, to ensure no significant loss in detection occurs between servicing by support staff. The recommended voltage result from the detector check should be within 100V of the current setting prior to generating fingerprint data.**

Guidelines

- A detector check should be performed monthly to monitor changes over time more closely, with the resulting recommended voltage used to maintain the detector voltage within 75V of the current setting.

Sample Preparation

Various sample preparations are outlined by the manufacturer and other sample preparation methodologies are extensively discussed in the literature. These standards do not reflect the support of a particular set of preparation methods, as data resulting from any preparation method is valid for use with fingerprint generation if the other standards are met. Rather, it is required to include information regarding the preparation used in the metadata describing the resultant fingerprint to ensure proper curation of shared library data.

Standard – Source material for generating reference spectra.

- **All specimens used to create reference spectra must be vouchered specimens stored in a reference collection. The voucher specimen and any genomic sequence data generated to confirm its identity must be clearly associated with the reference spectra via a voucher or collection number following the naming conventions specified under the Data Handling and Sharing section.**

Laboratory environment

The operating environment of the MALDI-ToF MS is also important, and manufacturer guidelines should be followed. Upkeep of the immediate laboratory environment as well as exterior instrument cleanliness is expected. This includes operation of the instrument within the recommended temperature ranges of 16°C - 30°C with the relative humidity between 20% - 75% non-condensing to prevent potential vacuum chamber, laser operation, or electronics issues.

Preparation of sample spots also requires consideration of the environment in which they are prepared. To ensure proper drying of the matrix used, it is important to consider the humidity levels of the room in which the drying sample is kept. Excessive drying times for the matrix can lead to degradation of the associated sample, and incomplete drying can lead to a loss in signal intensity, poor instrument sensitivity, and inconsistent results.

Standard – humidity and temperature

- **The instrument should be located in a laboratory area maintained within the recommended temperature ranges of 16°C - 30°C with a relative humidity between 20% - 75% (non-condensing) to prevent potential vacuum chamber/vacuum pump, laser operation, and/or electronics issues.**

Guidelines

- It is recommended that samples that will be used to generate fingerprint data are dried in an environment with lower than 70% humidity. Lower humidity correlates strongly with better sample and matrix drying and improved detection and consistency. This recommendation specifically includes the BTS used to calibrate and control quality during fingerprint data generation runs.

Growth media

Growth media for bacterial or fungal specimens that are sampled for fingerprinting with the MALDI-ToF MS should be noted in the metadata of the generated fingerprint. Certain growth media, such as blood agar or high-salt content media, are not recommended for use with MALDI-ToF MS, though there is nothing preventing generation of fingerprints from specimens cultured on those media. Rather, use of such media for specimen growth that is sampled for fingerprint generation may limit the fingerprint's identification ability by affecting its specificity. However, as long as the data quality standards outlined in this report are maintained, such fingerprints remain valid for both use and sharing.

Guidelines – growth media

- It is recommended that specimens cultured for sampling for fingerprint generation are cultured on low-salt content, virtually universal media, such as nutrient agar. This allows for a useful standard approach to be taken for specimens that are intended for MALDI-ToF MS diagnostics following fingerprint library generation.

Data Collection/Spectrum Generation

Calibration

In alignment with manufacturer recommendations, it is required to perform a calibration during the generation of spectra that are intended for use in a reference spectrum. Qualification of this timeframe *during,* is that if the data is collected with an automated data collection analysis (i.e., using the Bruker Compass software), a BTS must be properly labelled as a BTS control during the run set-up and analysed during the run for calibration and QC purposes. If the data is collected manually (i.e., using the Bruker FlexControl software), a

calibration must be manually performed on a spot prepared using BTS and the data for that spot must be collected within the sequence of manual data collections before being checked manually, as described by the manufacturer's guidelines regarding the production and development of MSPs (fingerprints).

Standards – calibration during data collection

- **A calibration, either automatic or manual, must be performed immediately prior to collection of fingerprint data.**
- **Mass shifts for each calibration peak should be ≤300 ppm. This is automatically assessed during automated calibration, and the calibration is rejected if any peak is ≥300 ppm from the expected position (using default Auto Calibration parameters). A manual calibration requires checking to ensure calibrant peaks fall within acceptable m/z precision constraints.**
- **Variation for calibration constants C_0, C_1, and C_2 should be within 3 standard deviations of their means for this calibration for fingerprint data to be considered valid (see MALDI-ToF Mass Spectrometer Maintenance and Health: Calibration/QC records for details on the requirements regarding recording these variables).**

Guidelines

- Variation for the calibration constants C_0, C_1, and C_2, should be within 2 standard deviations of their means.

Quality Control

Manufacturer recommendations regarding fingerprint (MSP) development includes a simplified quality control measure involving measuring the BTS spot used to initially calibrate the instrument and then manually analysing the results to ensure they match the expected calibration peaks. While sufficient to ensure a single known point is measured, this is not the same rigor applied to the automated QC check performed by the instrument to BTS prepared spots. For that reason, it is recommended to use automated data collection procedures to collect fingerprint data in conjunction with a quality control sample of BTS. In either case, the matching *E. coli* score for the BTS spot should be recorded (see MALDI-ToF Mass Spectrometer Maintenance and Health: Calibration/QC records).

Records for the BTS spot should account for the average score of the *E. coli* match, as well as the standard deviation. This enables the laboratory to ensure that any matching score variations are detectable. Matching scores below expected values can indicate a poor calibration and may invalidate any corresponding fingerprint data.

Standards – quality control during data collection

- **A BTS quality control must be analysed for every fingerprint data collection.**
- **Variation for a BTS matching *E. coli* score should be within 3 standard deviations of the trending mean for any data collection that is intended to contain reference spectra data.**

- **The BTS quality control must be analysed with the same calibration parameters as the fingerprint data collection (do not re-calibrate between QC and fingerprint data collection).**

Guidelines

- Use of automated data collection and its quality control validation is recommended to further ensure valid fingerprint results.
- Whenever possible, BTS *E. coli* matching scores should be checked to ensure variation is within 2 standard deviations of the mean.
- It is recommended that the prepared BTS generates *E. coli* matching scores of ≥2.20 to meet routine quality requirements similar to the ones the instrument is expected to perform under following installation and servicing by support staff (scores of ≥2.30 for a 1/10 dilution of BTS with a clean instrument).

Spectrum Assessment

Compiled assessment

The assessment of the spectra that constitute a fingerprint should consider the *Mass precision assessment*, the *Quantitative assessment*, and the *Qualitative assessment.* Following *Quantitative* and *Qualitative* assessment spectra constituting a fingerprint are considered of sufficientquality for diagnostic use if at least one is scored *High* and neither score *Low*, and is considered High quality if both assessments are scored *High*. See the following sections for details regarding the various assessments.

Standards – assessment compiling

- **If the *Mass precision assessment* fails for any spectra, those spectra must be rejected from inclusion in the final fingerprint. There must be enough spectra that pass the *Mass precision assessment* to generate a fingerprint with sufficient data collection quality for a valid fingerprint to be produced.**
- **The grade generated from the compiled sum of averages for each portion of the *Quantitative assessment* must be greater than *Low*.**
- **The subjective *Qualitative assessment* grade must be greater than *Low*.**

Guidelines

- Highest quality fingerprints are produced when both *Quantitative* and *Qualitative* grades are *High*
- The imperfect nature of the *Quantitative assessment* may cause an otherwise high-quality spectrum to be assessed poorly, or vice-versa. Therefore, a combination of *Medium* and *High* grades for the two assessments is sufficient to indicate high quality fingerprint data.
- A score of *Medium* for both assessments indicates a *Medium* quality fingerprint, which is not recommended for long-term use or sharing. It is sufficient as a short-term diagnostic tool that should be replaced with a higher quality fingerprint at the earliest opportunity.

Mass Precision

In alignment with manufacturer fingerprint development protocol, the data used to generate a fingerprint must be collectively assessed for precision. The minimum requirement for this assessment is the analysis of the peak variation between spectra for at least one peak per 1,000 m/z over the range of 2,000 to 10,000 m/z. When a range of 1,000 m/z does not have measurable peaks, no assessment of precision is needed for that range. Priority for choosing peaks to assess precision is given to the highest signal peaks, as they are the components of the spectra that will be key for diagnostically investigated samples. Variation exceeding 500 ppm invalidates a spectrum for use within a fingerprint, and should result in removal of that data from the used data set. Approximately 20 spectra collected using 240 laser shots each (default settings) is recommended as a minimum to ensure sufficient data density, good coverage of variability, and quality matching for diagnostics. Variation of this minimum is acceptable when paired with variation in total laser shot to ensure maintenance of data density and quality.

Standards – mass precision assessment

- **A precision assessment must be performed on multiple peaks within the range 2,000 m/z to 10,000 m/z for the spectra intended for use in fingerprint development. Each span of 1,000 m/z that features detected peaks should have at least 1 precision assessment performed on, at minimum, one peak within that range. Assessment priority should be given to the highest signal peak in each span of 1,000 m/z.**

Guidelines

- Manufacturer recommendations state that the use of 5 spectra for fingerprint generation is an absolute minimum[11].
- When using default data collection settings, it is recommended that at least 20 spectra that are sufficiently precise are used to generate a fingerprint.
- When alternate data collection settings (Peak Maldi) are used, the combined sum of laser shots for all spectra used to generate the fingerprint should exceed 5,000 laser shots, and a minimum of 5 spectra are recommended to be compiled to generate a fingerprint.

Quantitative Quality Assessment

Precision assessment alone is insufficient to determine spectrum quality. The aspects of a mass spectrum that contribute to whether it is of high-quality are varied, and, traditionally, any additional determination of quality was made qualitatively by assessing visually recognizable ques. Quantitative assessment is preferable, but it is non-trivial to capture the intricacies of a mass spectrum mathematically . An approximated approach to this was developed and includes a combination of a variety of simplified quantitative assessments.

These assessments are intended to be performed on data processed using FlexAnalysis, the software provided with the Bruker instrument, and its default settings for peak processing (using the default 100 peak detection maximum), baseline subtraction (multipolygon

baseline subtraction automatically applied twice with a 5 Dalton search window), and smoothing (Svaitsky-Golay with a frame size of 25 Daltons). Baseline subtraction and smoothing are to be used once each prior to peak detection. Peak information including m/z, S/N, quality factor (for known targets), resolution, intensity, and area, can be extracted from each spectrum following peak detection. These assessments use all detected values of m/z, S/N, intensity, and area from each spectrum to approximate different quality aspects quantitatively. These assessments are applied with the assumption that the CompassExplorer software provided with the Bruker instrument will be used to generate the fingerprints using the default settings to do so (up to 70 peaks within a fingerprint; see *Diagnostic data density: Number of diagnostically useful peaks* for more details).

Assessments are applied to each individual spectrum. Resulting scores for each individual spectrum assessment are averaged together, and, if desired, can then be multiplied by a weighting value. The subsequent result for each assessment category is then summed and a quality assessment for the compiled fingerprint is assigned based upon the final sum compared to the potential maximum score.

Standards – quantitative assessment compiling

- **The compiled sum of all averaged, and, if desired, weighted, assessments, is graded Low Quality, Medium Quality, or High Quality based upon the following criteria.**
    - **Sum ≤ 50% of the maximum score → Low Quality**
    - **50% of the maximum score < Sum < 75% of the maximum score → Medium Quality**
    - **Sum ≥ 75% of the maximum score → High Quality**

This grade can then be considered along with a qualitative assessment to generate a more comprehensively considered, less subjective, quality assessment for fingerprint development.

Noise detection: Low m/z noise

The lack of peaks detected within the range of 0-3000 m/z indicates either significant detection of peaks in the higher range of much higher intensity, or abundance of noise in that range obfuscating peak detection. The latter occurrence can be amplified by baseline subtraction, as a high baseline in the low m/z range for low signal spectra may result in exaggeration of low m/z noise following baseline subtraction. Furthermore, a significant amount of chemical noise contributed by the matrix occurs at lower m/z values[2-3]. Therefore, consideration of the number of successfully distinguished peaks in the low m/z range is a fairly good way to filter out low signal, low quality data.

Standards – quantitative assessment of low m/z noise

- **Count the number of detected peaks that are >3,000 m/z. Score each spectrum as follows.**
    - **Number of detected peaks = 0 → 0**
    - **0 < Number of detected peaks < 5 → 1**
    - **Number of detected peaks > 5 → 2**

- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Noise detection and diagnostic data density: Peak intensity comparison of low m/z and diagnostic m/z

This approximation is useful for mathematically assessing the impact of high intensity, non-diagnostically useful, low m/z peaks overshadowing the higher mass, diagnostically useful peaks due to the difference in their relative intensities[3]. Some diagnostically useful peaks that may be required for discrimination of closely related species or strains occur at low intensity, and therefore if they are not considered due to overshadowing by higher intensity peaks of lesser identification-related importance, there is a potential impact on the usefulness of the fingerprint[4]. The lack of peaks in the low mass range can also indicate significant noise, and this measurement simultaneously helps to assess this as well.

Standards – quantitative assessment of noise and data density using intensity

- **Sum the intensities for all m/z values <3,000 m/z (I_1) and sum the intensities for all m/z values between 3,000 m/z and 15,000 m/z (I_2). Score each spectrum as follows.**
  - **$I_1/I_2 < 0.04 \rightarrow 1$**
  - **$0.04 \leq I_1/I_2 \leq 0.16 \rightarrow 2$**
  - **$I_1/I_2 > 0.16 \rightarrow 0$**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Diagnostic data density: Approximation of peak broadening

Peak broadening is an issue due to how peaks are found in raw spectra. The centroid approach taken by default in the Bruker software relies on the ability to project a sharp peak from the highest point of a particular m/z down to the baseline and determines the shape of the curved peak sides using the restrictions of the instrument parameters. There can be a failure to find peaks between broadened near-neighbour peaks that overlap, leading to a potential loss of diagnostically important data[5]. As this data loss may not be replicated in all measurements of the same specimen, the potentially missing peaks from the spectra used to generate a fingerprint are relevant to fingerprint quality. This can be caused by instrument resolution limitations and may not be improvable through data collection or sample preparation changes[6]. Furthermore, peak broadening can impact spectrum specificity if it is too extreme, in the same way loss of peaks affects identification ability[7]. Consideration of this potential flaw is therefore useful for fingerprint evaluation. Its impact also needs to be assessed while considering the maximum detected peaks. Additionally, the Bruker software performs spectra compressions during default fingerprint creation, which helps reduce the impact of peak broadening if it is not extreme.

An approximation of peak base width is made by treating the mass peak as a triangle by using the area and intensity of the corresponding m/z value. Possible overlap between neighbouring approximated mass peak triangles is then checked. The signal intensity of these overlaps must then be assessed to determine if they could potentially be obscuring other m/z peaks. This is feasible since obscured peaks would only remain undetected if their intensity is less than half the lowest intensity of the overlapping peaks, so knowledge of the global minimum intensity and the intensity of the overlapping peaks is all that is needed. Potentially obscured peaks are then compared to the

number of diagnostically useful detected peaks to determine a percentage relating to the potentially missing diagnostic information. Therefore, contextually, these potential obscured peaks are only relevant if less than the maximum peaks were detected during processing (default settings maximum is 100 peaks detected).

Standards – quantitative assessment of diagnostic data density using peak broadening

- **If the number of detected peaks is at the maximum detectable during processing (100 by default), the ratio of potentially hidden diagnostically useful peaks is 0**
- **If there are potentially diagnostically useful hidden peaks the following steps must be taken**
    - **Approximations of peak base width must be made for each m/z peak by dividing the area by the intensity.**
    - **Potential overlap can then be calculated by using the m/z of each peak. To do so, subtract the base width from the m/z of a detected peak, then check if the resulting value. Then, consider the nearest, lower m/z peak added to the width of its own peak base width. The former value should be greater than the latter. If it is not, that an approximate overlap should be noted.**
    - **For each approximated overlap, consider the intensities of the two m/z peaks. If the lower intensity is higher than twice the minimum detected intensity for the whole spectrum, then count this overlap as potentially hiding a peak.**
    - **If the sum of the detected peaks and the total potentially hidden peaks is less than the maximum (100 peaks by default), then divide the number of potentially hidden peaks by the number of peaks within the number of peaks detected between 3,000 m/z and 15,000 m/z to generate a ratio describing the potential loss of diagnostic data.**
    - **If the sum of the detected peaks and the total potentially hidden peaks is greater than the maximum (100 peaks by default processing), then the ratio should be 100 minus the detected peaks divided by the number of peaks within the number of peaks detected between 3,000 m/z and 15,000 m/z to generate a ratio describing the potential loss of diagnostic data.**
- **The ratio is then used to generate a score in the following manner.**
    - **Ratio of potentially hidden peaks > 0.1 → 0**
    - **0 < Ratio of potentially hidden peaks < 0.1 → 1**
    - **Ratio of potentially hidden peaks = 0 → 2**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Diagnostic data density: Number of diagnostically useful peaks detected

This assessment considers the density of the data composing the fingerprint by judging detected peaks. Aspects that impact this assessment relate to the sensitivity of the instrument, the data collection method, the processing method, the fingerprint creation method, and the effectiveness of the extraction performed on the targeted specimen; therefore optimisation of this parameter requires consideration of each. Centroid peak detection with an expected resolution of 5 and a maximum peak number of 100 are the default peak detection settings of the FlexAnalysis software

used in the Bruker system. Therefore, only the highest intensity 100 peaks will be labelled during processing. Furthermore, by default, fingerprints created using default settings use only the highest 70 peaks that fall within 3,000 m/z to 15,000 m/z range. Combined, this means it is appropriate to assess the number of processed peaks that fall within the diagnostic range as a correlated approximation for the number of peaks expected within the fingerprint. As this value correlates to the number of captured ribosomal peaks, and the subsequent identification ability of the fingerprint, it is a useful measure of quality[7-8].

Standards – quantitative assessment of diagnostic data density using peak number

- **The number of detected peaks from 3,000 m/z to 15,000 m/z for each spectra should be quantified and scored.**
    - **Number of diagnostic detected peaks < 30 → 0**
    - **30 ≤ Number of diagnostic detected peaks < 70 → 1**
    - **Number of diagnostic detected peaks ≥ 70 → 2**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Signal clarity: Peak intensity

The overall signal intensity of the spectrum is a useful quality indicator[8], as low intensity signal spectra can generate unspecific fingerprints due to increased impact of noise, contamination, and low-intensity fragment or multimer peaks. Poor signal can also arise due to inhibited ionisation efficiency, caused by contamination, poor drying, laser issues, or environmental factors. As this assessment is intended to judge detection performed by the MALDI-ToF MS it considers the signal intensity for all peaks, including those outside the diagnostic range of 3,000 m/z to 15,000 m/z, and assess the sum of intensities against an arbitrary intensity value found to be consistent across mass spectra produced on Bruker instruments. Consequently, this assessment features wide scoring ranges to account for differing system results.

Standards – quantitative assessment of signal clarity using peak intensity

- **Sum the intensities (arbitrary units) for all detected peaks.**
    - **Peak intensity sum < 50,000 a.u. → 0**
    - **50,000 a.u. ≤ Peak intensity sum < 100,000 → 1**
    - **Peak intensity sum ≥ 100,000 → 2**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Signal clarity: S/N maximum in the diagnostic range

Another judge of signal clarity is an assessment of the S/N. This S/N assessment considers the maximum S/N value within the diagnostic range to help determine if the diagnostic range contains data of sufficient signal clarity for a high-quality fingerprint. The maximum S/N peak is also important to assess, as it generally contributes significantly to a fingerprint's diagnostic matching score.

Standards – quantitative assessment of signal clarity using maximum S/N

- **Determine the highest S/N value within the 3,000 m/z to 15,000 m/z range.**
    - **S/N < 20 → 0**

  - **20 ≤ S/N < 50 → 1**
  - **S/N ≥ 50 → 2**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Signal clarity: minimum relative intensity in the diagnostic range

Assessment of signal clarity within a spectrum also requires consideration of the intensity of the lowest detected peaks, yet the raw intensity in arbitrary units is not a descriptive value and cannot be easily compared across different instruments or collection settings. Therefore, the *relative* minimum intensity is the value worth considering when assessing the minimum peaks in a spectrum. The relationship between the quality of a spectrum's identification capability and its median relative intensity of ribosomal peaks has been explored, and correlating values were determined[8]. However, ribosomal peaks cannot be known *a priori* or easily determined for many MALDI-ToF spectra, therefore consideration of the universal minimum relative intensity can be substituted for the median ribosomal peak relative intensity if the value of the minimum is comparable to the desired value for the median. The values chosen for this assessment are based upon the literature values observed as optimal for spectral quality-ribosomal peak intensity median[8], though have been used to describe a desirable minimum relative intensity threshold.

Standards – quantitative assessment of signal clarity using the minimum relative intensity detected within the diagnostic range

- **Determine the minimum relative intensity value within the 3,000 m/z to 15,000 m/z range.**
  - **minimum relative intensity ≤ 0.1 → 0**
  - **0.1 < minimum relative intensity ≤ 0.12 → 1**
  - **minimum relative intensity > 0.12 → 2**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Signal clarity: S/N mean in the diagnostic range

Continued assessment of S/N within a spectrum is important, due to its impact on spectrum quality. This S/N assessment considers the mean S/N value within the diagnostic range. Similar to the consideration of the median, this assessment helps determine if most data in the diagnostic range have good signal clarity[9-10]. The values chosen for this assessment are based upon lower than default signal detection values, and therefore default data collection should produce high scores.

Standards – quantitative assessment of signal clarity using S/N mean

- **Determine the mean S/N value within the 3,000 m/z to 15,000 m/z range.**
  - **S/N ≤ 5 → 0**
  - **5 < S/N ≤ 10 → 1**
  - **S/N > 10 → 2**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Noise detection and data density: Peak acuity for noise and peak broadening detection

Noise and other undesirable signals often produce sharply defined peaks, commonly found below the diagnostic range in a MALDI-ToF MS spectrum. This assessment exploits this trend to

simultaneously judge a spectrum's noise as well as critique diagnostic peak broadening by comparing the area of peaks within the diagnostic range to peaks below the diagnostic range. Peak acuity, an approximation of its sharpness using the area of the peak divided by the intensity of the same peak, is used to for this assessment. The approximation assumes that, after processing, the diagnostic range should produce peaks only slightly broader than the lower m/z range, as centroiding and spline transformation generally lead to sharp apexes and broader bases for ion peaks when using a time-of-flight instrument with the potential for different vectors of kinetic energy[5], and penalizes either range for excessive broadening or sharpness due to those signifiers' relation to causing ambiguity through overlap and resulting from noise or poor signal, respectively.

Standards – quantitative assessment of noise and data density using peak width

- **Determine the area for peaks detected in the range of 2,000-2,999 m/z (A_1), the intensities for peaks in the range of 2,000-2,999 m/z (I_1), the area for peaks detected in the range of 3,000-15,000 m/z (A_2), and intensities for peaks in the range of 3,000-15,000 m/z (I_2). For each peak, find the value for the area divided by the intensity. Determine the average A_1/I_1 ratio and the average A_2/I_2 ratio and use these values to determine the spectrum score.**
  - **Avg(A_1/I_1) > [Avg(A_2/I_2)]/2 → 2**
  - **Avg(A_1/I_1) < [Avg(A_2/I_2)]/2 → 0**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Signal clarity: S/N mean of significant clarity peaks

Continued assessment of S/N within a spectrum is important, due to its impact on spectrum quality. This S/N assessment considers the mean of the S/N values of the top 20% S/N peaks in the spectrum. As these top peaks make up an important part of the fingerprint, their quality must meet minimum standards to ensure that the resultant fingerprint used markers that are sufficiently above the noise. An important difference in this measure as compared to using the overall mean is that it is not as easily affected by the minimum S/N value used during peak processing. , It better reflects the overall quality of the clearly defined peaks in the spectrum, and this is reflected in similar use of a range of the highest S/N peaks to measure quality in the literature[10]. The values chosen for this assessment are based upon consideration of potential minimum S/N values for intense peaks, and it is expected to be a lenient measure for high clarity spectra.

Standards – quantitative assessment of signal clarity using S/N mean of significant clarity peaks

- **Determine the mean S/N value of the top 20% S/N peaks in the spectrum.**
  - **S/N < 10 → 0**
  - **10 ≤ S/N < 20 → 1**
  - **S/N ≥ 20 → 2**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Signal reproducibility: reproducibility of fingerprint peaks

This assessment follows a similar criteria as the manufacturer recommended mass precision check, as outlined in the mass precision section of the spectrum assessment. However, this assessment

differs in that it considers the percentage of peaks that are reproduced across the processing range, using a 1-1 basis for each spectrum, rather than all the spectra considered against each other simultaneously. The intent is to capture the average amount of reproducibility between spectra as a percentage of the total number of detected peaks. Furthermore, the precision of the match should remain based upon the m/z ppm difference desired during the assessment, which, following the default manufacturer recommendation, is 500 ppm. This value has a notable impact on the resultant fingerprint and its identification quality, as peaks are weighted during identification matching by how well conserved they are across the spectra making up the fingerprint. Therefore, higher levels of reproducibility improve the consistency of the scoring. However, there is complexity around this value thanks to the lack of homology in the processing-to-fingerprinting assessment when using manufacturer provided programs FlexAnalysis and CompassExplorer. CompassExplorer, by default, uses 70 peaks from the 3,000-15,000 m/z range in a fingerprint, and in FlexAnalysis, the 100 highest intensity peaks from the 2,000-20,000 m/z range are labelled and described during default-setting processing. Therefore, while only 70% of potentially labelled peaks need to be reproduced in a spectrum with at least 100 peaks detected, they may not be in the desired range. Despite this imperfect homology, it is reasonable to assume that if at least 70% of FlexAnalysis processed peaks are reproduced, a similar amount would still be reproduced when the considered range is diminished in size. Furthermore, literature suggests that higher levels of reproducibility correlate with higher spectrum quality[4,8,9,11], and that this approach is sufficient to measure that aspect of the fingerprint. The values chosen for this assessment are based upon consideration of maximum number of peaks used in a fingerprint with default settings, and the similar literature outlined values for quality consideration.

Standards – quantitative assessment of peak reproducibility

- **Determine the number of peaks that are reproduced, within accepted m/z precision, across 2 spectra while judging the second spectrum against the first. Consequently, the very first spectrum in a fingerprint is not graded in this way.**
    - **Percent of reproduced peaks < 70%; and; Number of peaks < 100 → 0**
    - **Percent of reproduced peaks < 70%; and; Number of peaks ≥ 100 → 1**
    - **Percent of reproduced peaks ≥ 70% → 2**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Signal clarity: Inter-spectrum S/N conservation

Continued assessment of S/N is important, due to its impact on spectrum quality. This S/N assessment considers standard deviation of the S/N values for a selection of up to three peaks found in the spectra. The three peaks are to be found at different ranges of the m/z axis to enable better coverage and assessment of quality consistency. Functionally, this means three peaks are chosen from the first spectrum. More specifically, the highest peak in a specified m/z range is found, and this process is repeated across three ranges. After then collecting the S/N for those same peaks from each spectrum under consideration, the standard deviation of the S/N for each of the three peaks is found. This standard deviation is then normalised using the S/N mean for each of the three chosen peaks, and the resulting standard deviation percentage is then universally assessable. This enables a quantitative check of S/N consistency, similar to a method used in the literature to effectively assess the impact on result quality caused by changes to voltage and laser settings[12]. The values chosen for this assessment are based upon consideration of potential minimum S/N values for intense peaks, and it is expected to be a lenient measure for high clarity spectra.

Standards – quantitative assessment of S/N conservation between spectra

- **Determine the standard deviation of S/N values across all spectra for a set of peaks in a specified m/z range. These S/N values refer to the highest peak in three different ranges (3000-5500 m/z, 5500-8000 m/z, and 8000-10500 m/z). This set of up to three standard deviations are then normalised into a percentage value individually by dividing them each by the mean S/N value for that particular peak. The average between these values is then found and used in the following assessment.**
    - **S/N Standard deviation percentage average < 10 → 0**
    - **10 ≤ S/N Standard deviation percentage average < 20 → 1**
    - **S/N Standard deviation percentage average ≥ 20 → 2**
- **Average the scores for each spectrum together. Add this value to the compiled quantitative total.**

Semi-qualitative Quality Assessment

Mass spectrum assessment conventionally contains a qualitative component, as the varied aspects of a mass spectrum are difficult to completely assess quantitatively. This introduces a subjective element to spectrum quality assessment that may vary significantly between assessors. To enhance reproducibility, using a semi-quantitative approach that utilises a collection of questions for which the answers follow a 5-point Likert-scale grading scheme is recommended. Furthermore, the potential complexity of mass spectra necessitates that the visual, qualitative component must be weighted as equivalent to the entire quantitative component to ensure that poor quality spectra are properly filtered from high scores. Combined with the manufacturer recommended precision assessment and the quantitative quality assessments already outlined, the overall objectivity of spectrum quality assessment for fingerprint development should be improved..

Standards – qualitative assessment compiling

- **The compiled sum of all 5-point Likert-scale visual spectral assessments of the 2,000-20,000 m/z range is graded Low Quality, Medium Quality, or High Quality based upon the following criteria.**
    - **Sum ≤ 50% of the maximum score → Low Quality**
    - **50% of the maximum score < Sum < 80% of the maximum score → Medium Quality**
    - **Sum ≥ 80% of the maximum score → High Quality**

This grade can then be considered along with a quantitative assessment to generate a more comprehensively considered, less subjective, quality assessment for fingerprint development.

Qualitative assessment: Peak signal clarity

This visual assessment applies to the clarity of the ion peak signal and the noise in the spectrum. The fingerprint assessor should judge whether the peaks in the combined spectra are significantly above the noise or if the ion signal intensities are comparable to the noise, and rate this on a scale of 1 to 5, with 1 corresponding to the combined spectra featuring mostly low-signal ions and/or high noise, and 5 corresponding to low noise with peaks having visually obvious signal clarity.

Standards – qualitative assessment of peak signal clarity

- **Determine the peak signal clarity.**
  - **Low peak signal clarity or high noise → 1**
  - **Low noise and significantly clear peaks → 5**

Qualitative assessment: Peak sharpness

This visual assessment applies to the width and proximity of the ion peaks in the spectrum. The fingerprint assessor should judge whether the peaks in the combined spectra are sharp and well defined or if they are broadened and/or poorly resolved from each other, and rate this on a scale of 1 to 5, with 1 corresponding to the combined spectra featuring significant peak broadening, and 5 corresponding to well resolved peaks.

Standards – qualitative assessment of peak sharpness

- **Determine the peak sharpness.**
  - **Low peak sharpness → 1**
  - **Well-resolved, sharp peaks → 5**

Qualitative assessment: Peak distribution

This visual assessment applies to the distribution of the ion peak signals in the spectrum. The fingerprint assessor should judge whether the peaks in the combined spectra are well distributed across the diagnostic range (3,000-15,000 m/z) or if there is a poor distribution of peaks, and rate this on a scale of 1 to 5, with 1 corresponding to a poor distribution of peaks in the diagnostic range, and 5 corresponding to a good distribution of many peaks.

Standards – qualitative assessment of peak distribution

- **Determine the peak distribution.**
  - **Poor peak distribution → 1**
  - **Good distribution of many peaks → 5**

Qualitative assessment: Number of clear peaks in the diagnostic range

This visual assessment applies to the number of peaks detected in the spectrum within the diagnostic range (3,000-15,000 m/z). The fingerprint assessor should judge whether the number peaks in the combined spectra corresponds to a useful amount, and rate this on a scale of 1 to 5, with 1 corresponding to the combined spectra featuring few peaks in the diagnostic range, and 5 corresponding to many peaks being present and detectable.

Standards – qualitative assessment of the number of peaks in the diagnostic range

- **Assess the peak number in the diagnostic range.**
  - **Few diagnostic peaks → 1**
  - **Many diagnostic peaks → 5**

Qualitative assessment: Low m/z noise

This visual assessment applies to the low m/z noise (<3,000 m/z) in the spectrum. The fingerprint assessor should judge whether there is significant amounts of low m/z noise, which can be representative of poor signal or a high baseline, and rate this on a scale of 1 to 5, with 1

corresponding to the significant amounts of low m/z noise, and 5 corresponding to no visible noise in the low m/z range.

Standards – qualitative assessment of low m/z noise

- **Determine the amount of low m/z noise.**
  - **Significant amounts of low m/z noise → 1**
  - **No significant low m/z noise → 5**

Fingerprint Post-Production Assessment

Following production of a fingerprint, additional assessments should be performed to determine its functionality in the context of its diagnostic use. This includes assessment of its specificity for targets in its category of diagnostics as well as its diagnostic performance.

Standard – specificity assessment

- **An assessment of the fingerprint specificity should be performed following fingerprint production. Perform this assessment by attempting identification of the completed fingerprint using the supplemented library in CompassExplorer. Matching score criteria should fit the targeted specimen type in regard to near species and genus cut-off matching scores.**

Guidelines

- Providing available material, diagnostic performance should be tested using different specimens of the same species to determine diagnostic matching quality.
- If possible, inter-laboratory verification of the fingerprint should be assessed by sharing the fingerprint with another MALDI-ToF MS equipped laboratory and testing the same species of target as well as related targets.

Data Handling and Sharing

Fingerprint naming conventions

Universality in naming conventions allows simple sharing of pertinent development information without the need for metadata searching, enabling superior statistical analysis, while permitting use across laboratories with different degrees confidentiality or information management system requirements. Naming convention by needs must vary based on the taxonomic kingdom of the sample used to generate the fingerprint, to enable capture of key descriptors in the fingerprint name.

Standard

- **A naming convention adopted by a laboratory generating custom MALDI-ToF MS fingerprints must remain consistent and include certain key descriptors. For all sample taxonomic kingdoms, the required information presented in the name includes genus, species, subspecies, and a traceable registration code (potentially**

the institution name paired with a LIMS or sample voucher number) that can be used by the laboratory to access the fingerprinted sample's records.

Guidelines

- Bacteria and fungi recommended naming convention includes the following information in the presented order:

Genus	Species	Subspecies	Kingdom	Preparation method	host	Confirmation method	Institution	Accession no.	MALDI

Example: Neofusicoccum parvum FU full Clivia_miniata molec DAFF-BISS-385533_DNA130 MALDI

- Animalia naming convention includes the following information in the presented order:

Genus	Species	Subspecies	Kingdom	Life stage	Body part	sex	caste	host	Confirmation method	Institution	Accession no.	MALDI

Example: Amarygmus tristis AN adult midleg_tibia_tarsus molec DAFF-BISS-449144 MALDI

- Plantae naming convention includes the following information in the presented order:

Genus	Species	Subspecies	Kingdom	Life stage	Part	Confirmation method	Institution	Accession no.	MALDI

Example: Oxalis sp PL seed whole morph DAFF-PDI-363619 MALDI

- Archaea, Protozoa, and Chromista naming convention includes the following information in the presented order:

Genus	Species	Subspecies	Kingdom	Host or microhabitat	Confirmation method	Institution	Accession no.	MALDI

- For all presented conventions, the shaded cells indicate optional data, and, when this information is not available or known, it should be removed from the compiled name.
- *Kingdom* is shortened to the first 2 letters of the kingdom name with both letters capitalized.
- *Confirmation method* is shortened to the first 5 letters of either molecular (molec) or morphological (morph).
- Institution and accession no. should be separated by a dash rather than a space to indicate they are linked. Accession no. refers specifically to the manner used to trace and track the sample within the laboratory making the fingerprint. This may correspond to a voucher number, traceable sample laboratory number, or other alpha-numeric sample registering code. The institution must retain access to relevant sample information, including identification confirmation records, sample source, etc., though this number, and be able confirm correctness of recorded fingerprint information through this record system.

Resources

MBT Compass User Manual (Sept. 2014). Revision D (April 2018). Bruker Daltonik GmbH.

MBT Explorer Module User Manual (Feb. 2016). Revision A (May 2016). Bruker Daltonik GmbH.

References

[1] Karkalousos, Petros & Evangelopoulos, Angelos. The History of Statistical Quality Control in Clinical Chemistry and Haematology (1950 – 2010). *International journal of biomedical science* **2015**, *4*, 1-11.

[2] Krutchinsky, A. N.; Chait, B. T. On the Nature of the Chemical Noise in MALDI Mass Spectra. *Journal of the American Society for Mass Spectrometry* **2002**, *13* (2), 129–134. https://doi.org/10.1016/S1044-0305(01)00336-1.

[3] Vestal, M. L. Modern MALDI Time-of-Flight Mass Spectrometry. *Journal of Mass Spectrometry* **2009**, *44* (3), 303–317. https://doi.org/10.1002/jms.1537.

[4] Sindt, N. M.; Robison, F.; Brick, M. A.; Schwartz, H. F.; Heuberger, A. L.; Prenni, J. E. MALDI-TOF-MS with PLS Modeling Enables Strain Typing of the Bacterial Plant Pathogen *Xanthomonas Axonopodis*. *Journal of the American Society for Mass Spectrometry* **2017**, *29* (2), 413–421. https://doi.org/10.1007/s13361-017-1839-0.

[5] Urban, J.; Afseth, N. K.; Štys, D. Fundamental Definitions and Confusions in Mass Spectrometry about Mass Assignment, Centroiding and Resolution. *TrAC Trends in Analytical Chemistry* **2014**, *53*, 126–136. https://doi.org/10.1016/j.trac.2013.07.010.

[6] Tsypin, M.; Asmellash, S.; Meyer, K.; Touchet, B.; Roder, H. Extending the Information Content of the MALDI Analysis of Biological Fluids via Multi-Million Shot Analysis. *PLOS ONE* **2019**, *14* (12), e0226012. https://doi.org/10.1371/journal.pone.0226012.

[7] Müller, P.; Pflüger, V.; Wittwer, M.; Ziegler, D.; Chandre, F.; Simard, F.; Lengeler, C. Identification of Cryptic Anopheles Mosquito Species by Molecular Protein Profiling. *PLoS ONE* **2013**, *8* (2), e57486. https://doi.org/10.1371/journal.pone.0057486.

[8] Cuénod, A.; Foucault, F.; Pflüger, V.; Egli, A. Factors Associated with MALDI-TOF Mass Spectral Quality of Species Identification in Clinical Routine Diagnostics. *Frontiers in Cellular and Infection Microbiology* **2021**, *11*. https://doi.org/10.3389/fcimb.2021.646648.

[9] Mitchell, M. W.; Mali, S.; King, C. C.; Bark, S. J. Enhancing MALDI Time-Of-Flight Mass Spectrometer Performance through Spectrum Averaging. *PLOS ONE* **2015**, *10* (3), e0120932–e0120932. https://doi.org/10.1371/journal.pone.0120932.

[10] Drevinek, M.; Dresler, J.; Klimentova, J.; Pisa, L.; Hubalek, M. Evaluation of Sample Preparation Methods for MALDI-TOF MS Identification of Highly Dangerous Bacteria. *Letters in Applied Microbiology* **2012**, *55* (1), 40–46. https://doi.org/10.1111/j.1472-765x.2012.03255.x.

[11] Seuylemezian, A.; Aronson, H. S.; Tan, J.; Lin, M.; Schubert, W.; Vaishampayan, P. Development of a Custom MALDI-TOF MS Database for Species-Level Identification of Bacterial Isolates Collected from Spacecraft and Associated Surfaces. *Frontiers in Microbiology* **2018**, *9*. https://doi.org/10.3389/fmicb.2018.00780.

[12] Wetzel, S.; Guttman, C. M.; Flynn, K. M.; Filliben, J. J. Significant Parameters in the Optimization of MALDI-TOF-MS for Synthetic Polymers. *Journal of the American Society for Mass Spectrometry* **2006**, *17* (2), 246–252. https://doi.org/10.1016/j.jasms.2005.11.007.