Evaluation of MALDI-TOF mass spectrometry for identification of *Mycobacterium* species found during environmental monitoring

Warren Crabb, Christine E. Farrance, Prasanna D. Khot

charles river

1 ABSTRACT

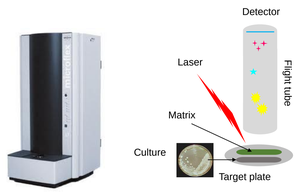
The genus *Mycobacterium* contains nearly 160 species and many are found in the natural environment in water and soil, which act as reservoirs for human and animal infection. Due to the likelihood of mycobacteria entering industrial water systems and being recovered during environmental monitoring activities, it is appropriate to review the identification of these species.

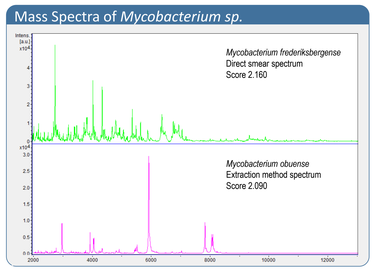
Recently, MALDI-TOF mass spectrometry has emerged as a leading option for identifying most bacteria and yeast based on abundant protein profiles. However, mycolic acids and complex lipids make the mycobacterial cell wall thicker, hydrophobic and waxy. These characteristics may pose a potential hurdle to accessing proteins using routine sample preparation methods for MALDI-TOF. We evaluated the identification rate for over 1500 *Mycobacterium* isolates, representing 43 unique species, groups, or complexes, using a reference library that included Bruker's latest Mycobacterium Library version 5.0. The identification rate was calculated using score thresholds validated at Charles River. Reference identification based on the 16S rRNA gene was used to evaluate accuracy of the MALDI result in a subset of the samples.

The identification rates by MALDI were 66% and 34% with and without Bruker's supplemental Mycobacterium library, respectively. These results show that MALDI has the promise to be an effective platform for identification of mycobacteria, but supplementing the library with a diversity of isolates is critical to improving performance. However, DNA sequencing is needed to identify the isolates which remain unidentified using MALDI.

2 STUDY COHORT

- 1540 isolates from the genus *Mycobacterium* were analyzed by comparing the current CRL MALDI library to Bruker's supplemental Mycobacterium Library version 5.0.
- All evaluated samples were submitted for MALDI processing between 2011 to 2018 by 215 customers in North America and Europe.
- Samples that originally failed to be identified by MALDI due to signal that was below threshold were processed by 16S rRNA gene sequencing.





3 LIBRARY EVALUATION

- The CRL MALDI Library contains 88 *Mycobacterium* strains representing 49 unique species. The Bruker supplemental Mycobacterium Library version 5.0 adds 912 strains representing 164 unique species.
- A total of 14 species group/complexes are described that are difficult to resolve by MALDI or sequencing.

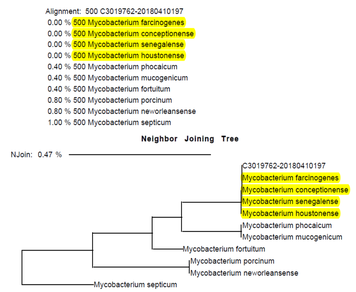
MALDI

Species of *conceptionense / farcinogenes / houstonense / senegalense* of the genus *Mycobacterium* have very similar patterns: Distinguishing their species is difficult.

Rank (Quality)	Matched Pattern	Score Value	NCBI Identifier
1 (++)	Mycobacterium senegalense DSM 43656T DSM b	1.89	1796
2 (++)	Mycobacterium senegalense DSM 43653 DSM b	1.85	1796
3 (++)	Mycobacterium senegalense DSM 43660 DSM b	1.83	1796
4 (++)	Mycobacterium senegalense DSM 43245 DSM b	1.8	1796
5 (++)	Mycobacterium fortuitum_complex T110 IMK b	1.79	1763
6 (++)	Mycobacterium senegalense 120307_10 CHL b	1.76	1796
7 (++)	Mycobacterium fortuitum ssp fortuitum DSM 46621T DSM b	1.76	144549
8 (++)	Mycobacterium senegalense DSM 43664 DSM b M	1.75	1796
9 (-)	Mycobacterium senegalense DSM 43294 DSM b	1.7	1796
10 (-)	Mycobacterium fortuitum BlutK LSM	1.7	1766

SEQUENCING

Mycobacterium conceptionense / farcinogenes / houstonense / senegalense are not resolvable with 16S rDNA sequencing and phylogenetic analysis



4 RESULTS

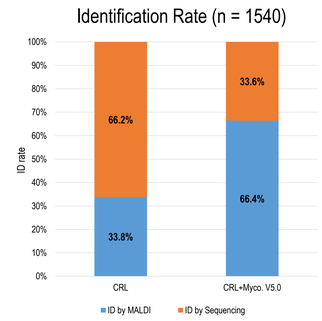


Figure 1. Comparison of the ID rate using the CRL library with the addition of the Bruker Mycobacterium 5.0 supplemental library. Inclusion of the supplemental library increased the MALDI identification rate by 32.6%

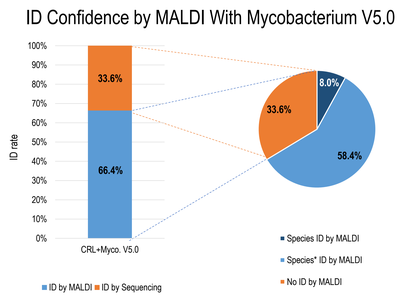


Figure 2. MALDI Identification confidence level results. 58.4% of the samples were identified to the group/complex level. 8% were resolved to the species level and the remaining 33.6% could not obtain an ID by MALDI. These samples were identified by sequencing the 16S rRNA gene.

5 RESULTS

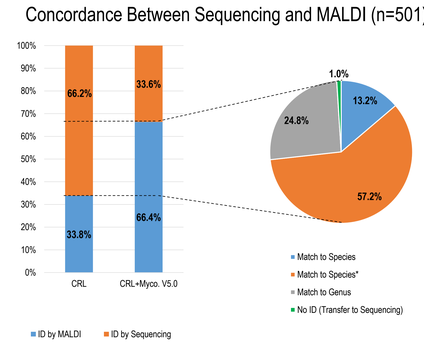


Figure 3. Concordance between newly identified MALDI samples and samples previously identified by sequencing. 57.2% samples match at the group/complex level and 13.2% matched to the species level. For 24.8% of samples only a genus-level sequencing ID was appropriate. A small fraction (1%) of samples provided a MALDI identification but did not meet CRL confidence-level criteria and require sequencing to provide an accurate identification.

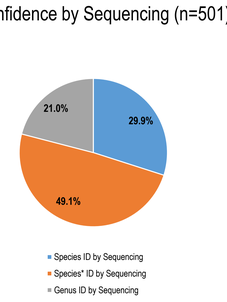


Figure 4. Confidence level breakdown for previously sequenced samples.

6 CONCLUSIONS

Bruker's Supplemental Library Mycobacterium version 5.0 has a significant operational impact on identification rate. However, backup identification method, such as rRNA gene sequencing, will have to continue to play a prominent role for this genus.

Rapidly growing Mycobacteria remain very difficult to speciate by MALDI. This is not unique as this same set of organisms is also difficult to speciate using 16S rRNA gene sequencing and a phylogenetic analysis – 50% of this study cohort were identified at the complex or group level. Further identification techniques will be required to separate these groups.

If a presumptive *Mycobacterium* identification is known, a modified extraction method may be beneficial in improving MALDI spectral generation.