

Operational impact of library development on the identification of organisms in the family *Bacillaceae* in pharmaceutical manufacturing using MALDI-TOF mass spectrometry backed up by sequencing

Christine E. Farrance, Sunhee Hong and Prasanna D. Khot

1 ABSTRACT

The family *Bacillaceae* comprises a wide phylogenetic diversity of approximately 70 genera. Members are commonly found in pharmaceutical manufacturing facility microflora. Most species form spores, thus, they are prevalent contaminants and resistant to disinfection. Even as MALDI-TOF-MS (MALDI) is becoming widely available as an identification method, there is little published regarding its efficacy for the *Bacillaceae*. Over 7 years, Charles River has expanded our MALDI library with nearly 400 entries that now represent 67.9% of the total entries for the family *Bacillaceae*. In this retrospective study, we analyzed the identifications of approximately 14,200 isolates spanning 20 genera from the *Bacillaceae* that were submitted for identification by MALDI. If the MALDI failed to provide an identification, the isolates were identified by 16S rDNA sequencing. Operational impact of using MALDI for identification was calculated for factors such as sequencing as a backup identification system, additional MALDI library development, sample preparation methods, and desired taxonomic resolution. In addition, impact of capturing strain diversity within the MALDI library and the option to discriminate closely related species using protein coding gene sequencing was reviewed. This study demonstrated the operational benefit of MALDI library development, backed up with sequence-based identification for accurate and rapid identification.

2 STUDY COHORT

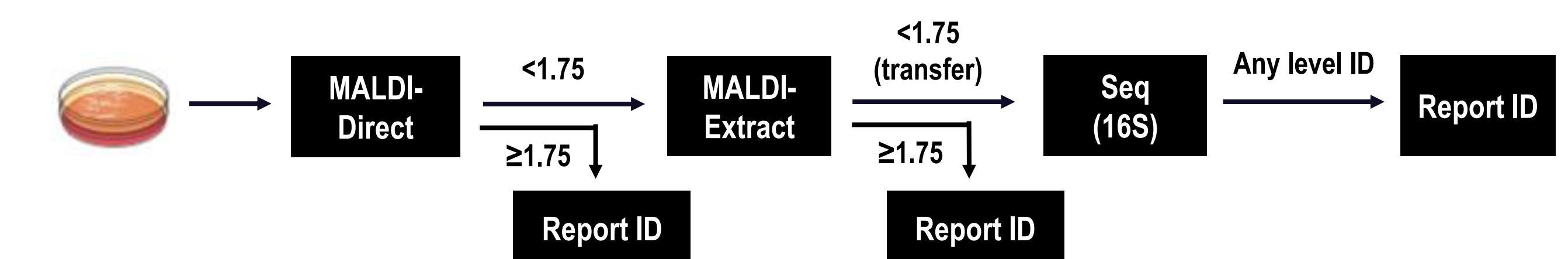
- Analyzed 14,181 isolates from family *Bacillaceae* submitted for MALDI backed by sequencing from 2011 to 2018 by 662 customers in North America and Europe
- These isolates represent 213 species or species groups of the family *Bacillaceae* spanning 20 genera

Top Six Frequently Identified Genera	Frequently Identified Species Within the Genera
<i>Bacillus</i>	<i>B. cereus</i> group, <i>B. licheniformis</i> , <i>B. circulans</i>
<i>Lysinibacillus</i>	<i>L. fusiformis</i> , <i>L. boronitolerans</i> , <i>L. massiliensis</i>
<i>Geobacillus</i>	<i>G. stearothermophilus</i> , <i>G. kaustophilus</i> , <i>G. thermodenitrificans</i>
<i>Oceanobacillus</i>	<i>O. profundus</i> , <i>O. sojiae</i> , <i>O. caeni</i>
<i>Psychrobacillus</i>	<i>P. insolitus</i> , <i>P. psychrodurans</i> , <i>P. psychrotolerans</i>
<i>Virgibacillus</i>	<i>V. halodenitrificans</i> , <i>V. pantothenticus</i> , <i>V. proomii</i>

3 STUDY DESIGN

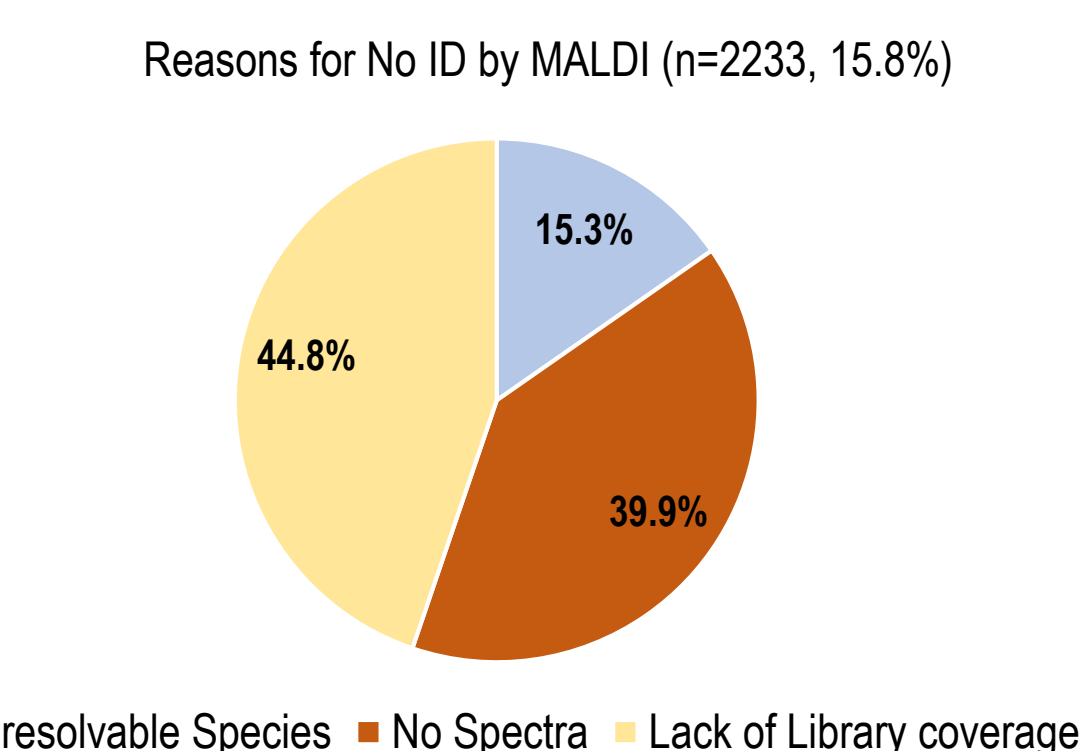
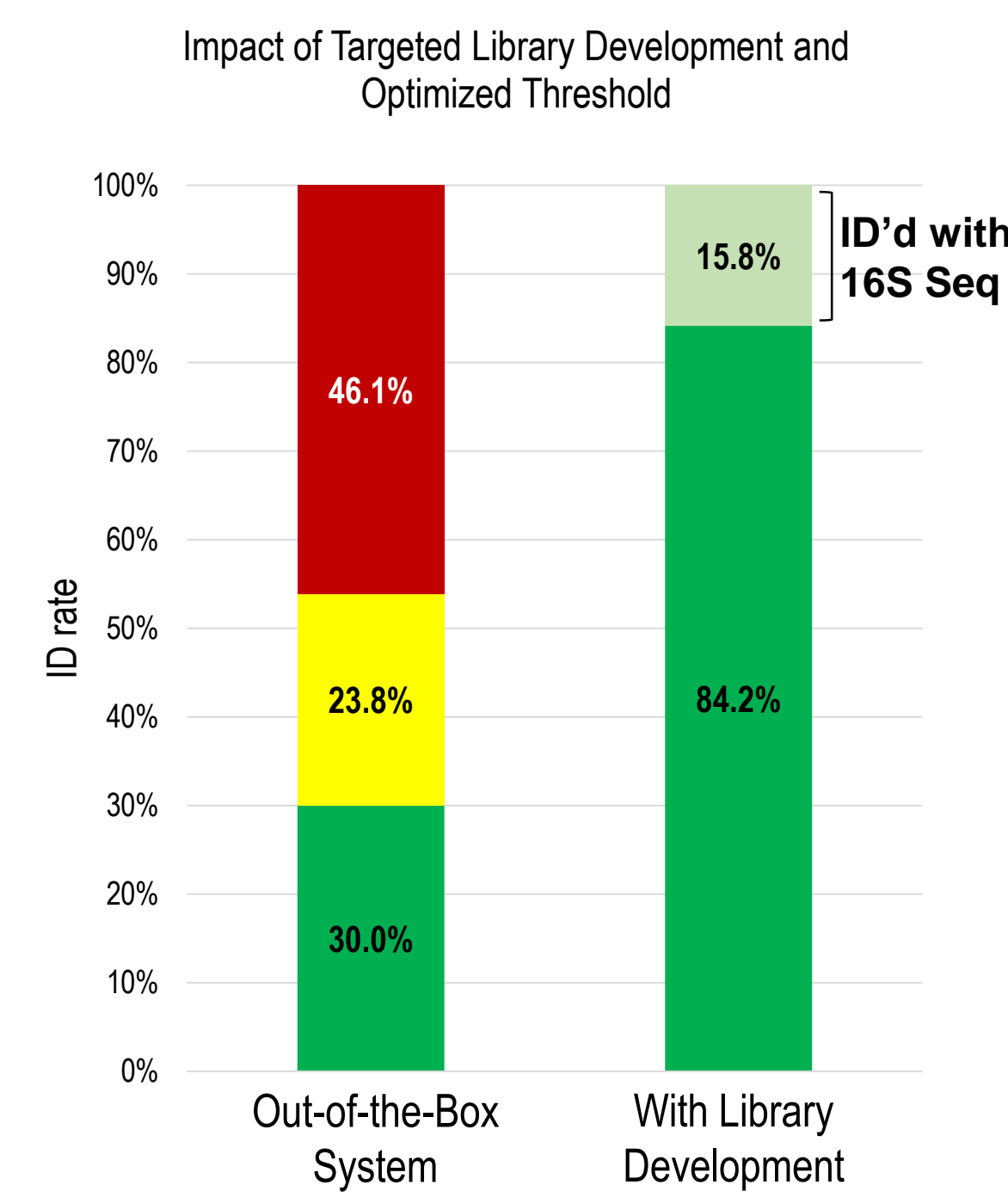
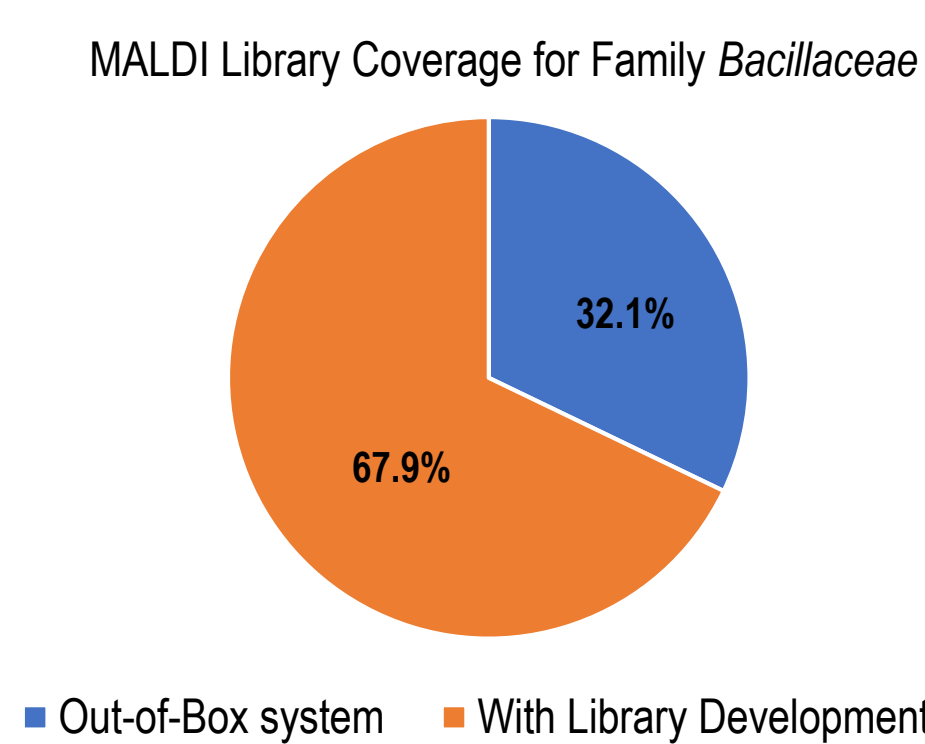
- Retrieved data from CRL's AccuPRO-ID® service, which is MALDI-TOF MS backed up by sequencing
- Retrieved ID results for each species until 200 samples were reached for frequently identified species or up to 100 for less frequently identified species
- Tracked reasons for No ID by MALDI (Transfer-to-Sequencing):
 - Due to top species scores too close (unresolvable to single species)
 - Due to No Match (lack of library coverage)
 - Due to No Spectra (suboptimal sample preparation)
- Calculated ID rate with and without the impact of targeted library development

Workflow of MALDI Backed By Sequencing (AccuPRO-ID Service)



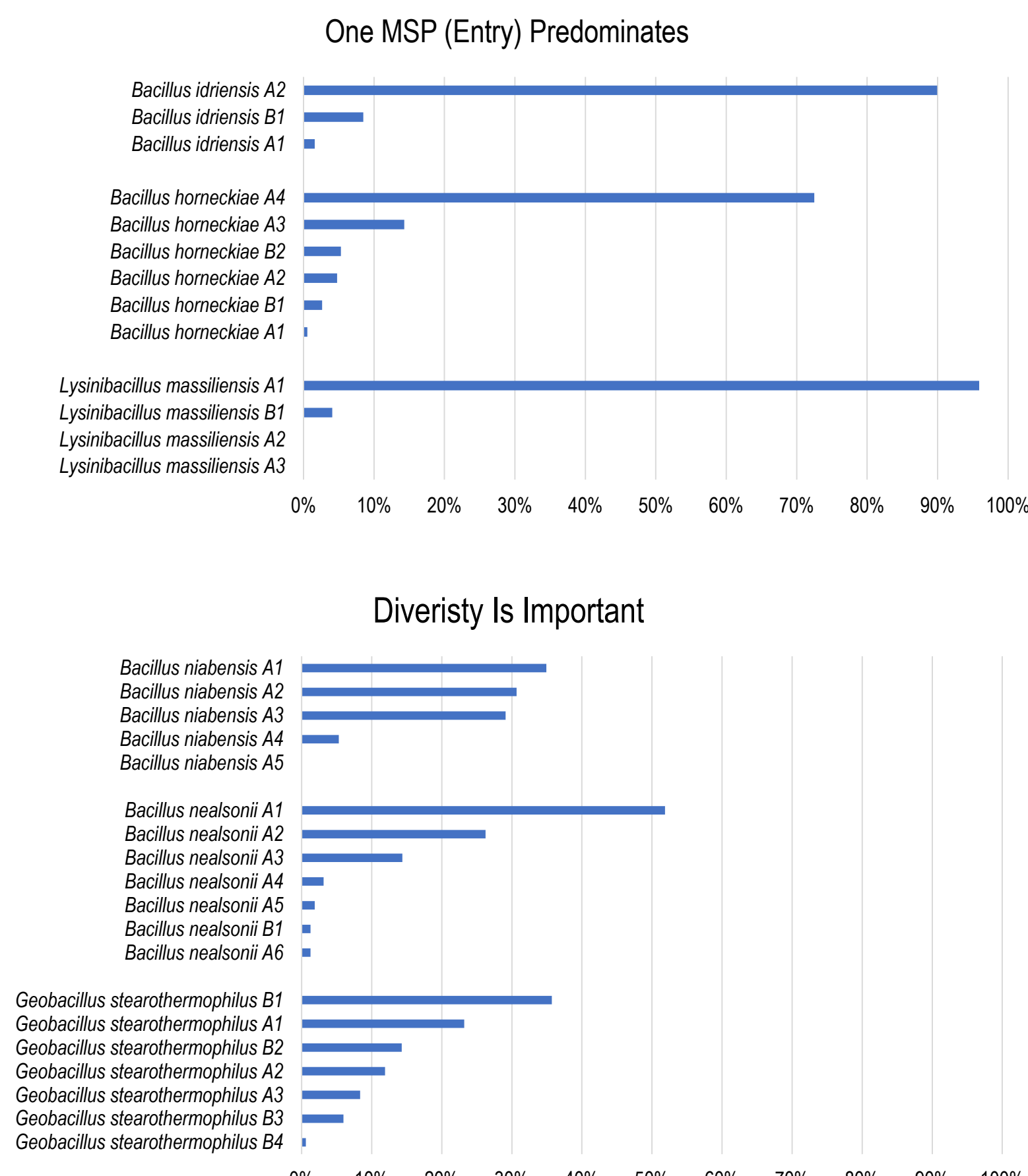
4 RESULTS

Impact of Targeted Library Development (n=14,181)

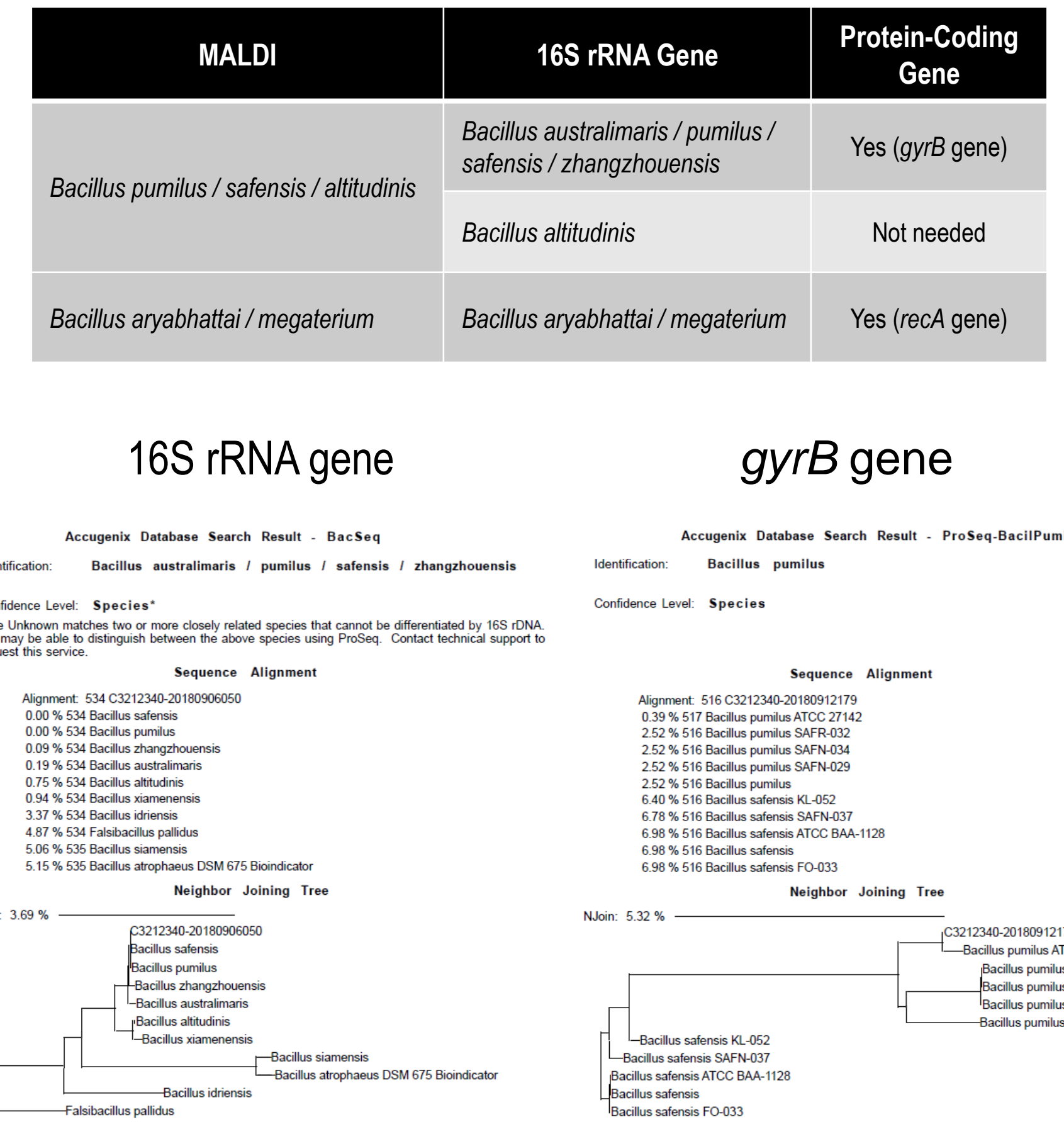


MALDI Score Key	Bruker	CRL
Species ID	≥2.0	≥1.75
Low confidence ID	≥1.7 and <2.0	N/A
No ID	<1.7	<1.75

Impact of Strain Diversity in the MALDI Library



Species Resolution



5 CONCLUSIONS

- Targeted library development is critical to improving ID rate
- Capturing strain diversity during library development is necessary
- Sequencing backup is very effective to ID isolates and resolve species MALDI cannot
- Specialized sample preparation may be needed for some species that do not generate spectra on MALDI by routine methods