Milking DNA

Researchers at TEAGASC have evaluated methods to improve DNA sequencing of the milk microbiome.

Milk is an important source of nutrition for both humans and animals. Human breast milk is highly beneficial to a child's development and milk from animals, particularly cows, is widely consumed across the globe. The study of the microbial communities found in milk is necessary from the perspective of both human and animal maternal health. It is also important to understand the impact of these communities on the safety and quality of milk used for consumption. High-throughput DNA sequencing approaches have been a valuable tool in this regard, providing information on milk microbiomes to reveal beneficial or harmful bacteria. Targeted amplicon sequencing, which mainly uses the 16S rRNA gene (common to all bacteria), has been adopted for the study of many diverse microbiomes, such as the human gut and soil. Shotgun metagenomic sequencing, which analyses the DNA of an entire sample, provides much greater insights regarding the microorganisms present, what they can do, and even allows for the generation of metagenome-assembled genomes (MAGs). These MAGs provide essentially complete genome information for key microorganisms identified in the sample, as well as revealing additional functional and safety properties associated with the microbial community (Bowers *et al.*, 2017). However, as shotgun metagenomic sequencing is an untargeted approach, DNA from the host (human or animal) cells present in the milk is also sequenced, which in the case of milk represents a considerable majority (up to 95 %) of the DNA present. This high proportion of host DNA results in wasted sequencing capacity (lots of host sequence information that is not of microbiological interest) and insufficient sequencing depth of the microbial DNA. To address this challenge, we evaluated different methods to either deplete host DNA or enrich microbial DNA using commercially available kits.

Both bovine and human milk samples were used for the study. Bovine milk samples were collected from farms across Ireland and human milk was collected from mothers in the Microbe Mom study, following ethical approval and with informed consent. Milk samples underwent several washing steps to remove the sample fat before DNA extraction and host depletion/microbial enrichment with three methods. The three methods evaluated are the DNeasy PowerSoil Pro kit (Qiagen), MolYsis complete5 kit (Molzym Gmbh & Co.), and NEBNext Microbiome DNA Enrichment Kit (New England Biolabs). A 10-strain mock community (consisting of 10 known microorganisms) was spiked into a milk sample as a positive control. Shotgun sequencing libraries were prepared from the subsequent DNA samples before sequencing on the Illumina NextSeq 500 platform at the Teagasc Sequencing Facility. Bioinformatic analysis on the shotgun metagenomic reads assigned both taxonomy (names of microorganisms) and genetic functional potential (what these microorganisms can do) of the milk microbiome.

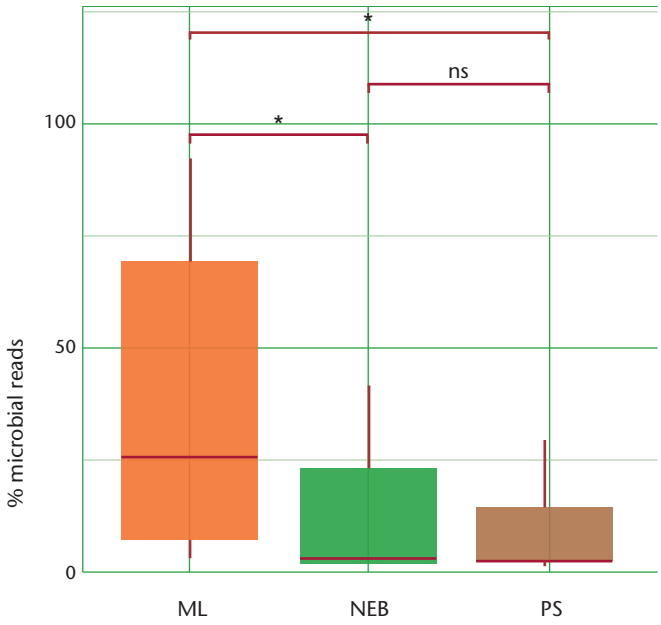

FIGURE 1: Comparison of mean microbial sequencing reads for all samples between the evaluated methods.

Results

We found that the MolYsis complete5 kit (ML kit) was efficient in depleting host DNA, enabling greater sequencing depth of microbial DNA compared to the other two kits evaluated (**Figure 1**). This method improved microbial reads by 20 %. Following bioinformatic analysis, we discovered that the choice of taxonomic classification tool had a

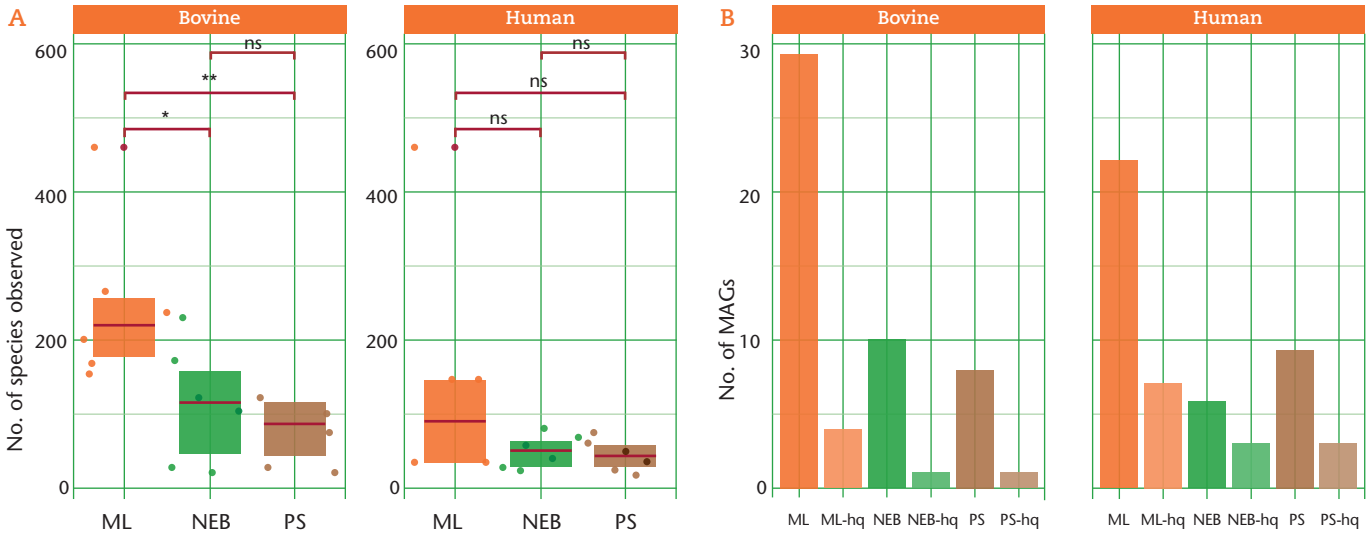

FIGURE 2: Better characterisation of the bovine and human milk microbiome was found for both observable species (a), and number of MAGs and high-quality MAGs (b) when the ML kit was used compared to the other two methods.

greater impact on the reported composition than the method used. The performance of taxonomic classification tools varied when milk samples containing the mock community were compared. Ultimately, one of these tools, Kraken2, was selected for further use as it performed best in terms of overall correct assignment and expected abundances of mock community DNA. The ML kit not only gave a significantly higher percentage of microbial reads, but the greater microbial sequencing depth enabled better characterisation of the milk microbiome after bioinformatics analysis. More unique bacterial species were detected, and more MAGs, and specifically high-quality MAGs, were recovered from the samples that used the ML kit than from the other two methods (**Figure 2**). Importantly, when comparing the community structure between methods, no biases were found.

Conclusion

Overall, this evaluation has addressed two important issues in metagenomic sequencing of the milk microbiome: specifically, poor microbial sequence depth and poor sequencing economics. The results show that the host depletion approach of the ML kit performed better than the enrichment or direct sequencing alternatives by providing the potential for deeper strain-level analysis without an observable bias. The improved sequencing of the milk microbiome that will be provided by this approach will be hugely beneficial in the agricultural, processing and clinical settings, as providing greater characterisation of the microbes present in milk samples can be used to inform food safety/quality practices and treatments.

Acknowledgement

This work was funded by the Irish Dairy Levy and Science Foundation Ireland (SFI) under grant numbers SFI/12/RC/2273 and 16/SP/3827, with co-funding by PrecisionBiotics Ltd. The authors acknowledge the contributions of Calum Walsh, Elaine Lawton and Fiona Crispie in the sequencing and bioinformatics analysis, and David Gleeson and Lizandra Paludetti in bovine milk collection. Additional thanks to the Microbe Mom team.

References

Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T., *et al.* (2017). 'Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea'. *Nature Biotechnology*, 35 (8): 725-731.

Yap, M., Feehily, C., Walsh, C.J., Fenelon, M., Murphy, E.F., McAuliffe, F.M., *et al.* (2020). 'Evaluation of methods for the reduction of contaminating host reads when performing shotgun metagenomic sequencing of the milk microbiome'. *Scientific Reports*, 10 (1): 1-11.

Authors

Min Yap
Walsh Scholar, Food Biosciences Department, Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork
Correspondence: min.yap@teagasc.ie

Conor Feehily
Postdoctoral Fellow, Food Biosciences Department, Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork

Paul Cotter
Head of Department, Senior Principal Research Officer, Food Biosciences Department, Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork