Proc. Assoc. Advmt. Anim. Breed. Genet. 26: 241-244

EXPLORING ADAPTIVE SEQUENCING FOR EFFICIENT PROFILING OF METHANOGENIC MICROBIOMES IN CATTLE ORAL SAMPLES AS A PROXY FOR METHANE PREDICTION AND MITIGATION STUDIES

Z. Chen[1], J. Silaban[2], E.M. Ross[1] and C.T. Ong[1]

[1] Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St. Lucia, QLD, 4072 Australia
[2] School of Chemistry and Molecular Biosciences, University of Queensland, St. Lucia, QLD, 4072 Australia

SUMMARY

Oxford Nanopore Technologies (ONT) adaptive sequencing enables real-time sequence enrichment or depletion during the sequencing run. However, the efficiency of different adaptive modes on microbiome analysis has not been well characterised. This study sequenced DNA from bovine saliva samples with two adaptive sequencing protocols (depletion and enrichment) and one regular protocol (adaptive function disabled). The depletion protocol with a bovine reference file efficiently reduced the bovine base number in the sequencing output. The enrichment protocol, despite using a methanogen reference file, resulted in the lowest archaeal species diversity. Greater base numbers of targeted methanogens were observed in the depleted samples compared to both the regular and enrichment protocols. The depletion protocol showed no notable effect on the relative abundances of selected archaeal genera. This study demonstrated the high efficiency of ONT depletion sequencing in minimising host contamination of saliva microbiome data, while effectively retaining the original microbial profile.

INTRODUCTION

The process of rumen fluid collection is invasive for animals. Because of the rumination process, oral swabbing has been considered a non-invasive alternative method for rumen microbial profiling. However, it also raises concerns regarding host genome contamination in the samples. Chemical and enzymatic host DNA depletion protocols have been widely used, but they also introduce biases to microbiome analysis (Ganda *et al.* 2021). In recent years, Oxford Nanopore Technologies (ONT) has introduced adaptive sequencing for the enrichment or depletion of sequences of interest and has shown potential applications in reducing host contamination (Ong *et al.* 2022). However, the efficiencies of different ONT adaptive strategies on microbiome characterisation require further evaluation. Therefore, this study hypothesised that ONT adaptive sequencing would significantly reduce host contamination. In the resulting data, we also tested the hypothesis that the relative abundances of archaeal genera would be affected by adaptive sampling.

MATERIALS AND METHODS

Sample collection. The bovine saliva samples of this study were collected under Animal Ethics Number AE000438. The animals used in this study were bulls housed at one of the commercial properties in Queensland, Australia. The Super•SAL™ device (Oasis Diagnostics® Corporation, USA) was used for bovine saliva collection following the manufacturer's instructions with minor modifications. Generally, the sponge was applied to gently swab the inner oral lining of a single bull until the indicator of the device turned fully red. Afterwards, the sponge was cut with sterile scissors and transferred into a 10 ml screw cap tube with 2 ml sterile 7:3 1xPBS-glycerol buffer, followed by ten inversions for mixing. After collection, samples were promptly placed on ice and transferred to -80°C for long-term storage. Samples from three animals (animals 7219, 7233, and 7234) were selected for testing.

DNA extraction and sequencing. A total of three biological replicates (animals 7219, 7233, and 7234) were used for DNA extraction in this study. The saliva was first squeezed from the sponge into a 10 ml screw cap tube by syringe. The squeezed liquid was centrifuged at 14,000 rpm for 5 min at 4℃, followed by removing the supernatant. The pellets were used for DNA extraction by QIAamp PowerFecal Pro DNA Kit (QIAGEN, Germany) according to the manufacturer's instructions. The DNA library was prepared by ONT Native Barcoding Kit 96 V14 (ONT, UK) following the manufacturer's protocol and subsequently loaded on a FLO-MIN114 (R10.4.1) Flow Cell. The sequencing was performed on the PromethION 2 Solo (ONT, UK). For samples sequenced with depletion protocol, a bovine reference file with *ARS-UCD1.2 Bos taurus* (GCA_002263795.2) and Brahman genomes (Ross 2019) was uploaded. Samples sequenced with enrichment mode utilised a reference file with 12 methanogen genomes. In addition, the adaptive function was disabled for samples under the regular protocol. Each adaptive strategy included three biological replicates, each with three technical replicates. The sequencing was terminated after all samples reached 300 Mb.

Bioinformatics and analysis. Sequencing data were base called using ONT Dorado v0.6.2 under the super accurate model with the adapter trimming option enabled. Reads shorter than 100 bp were removed using Nanofilt v.2.8.0 (De Coster *et al.* 2018). The quality-controlled data were mapped to the bovine reference file with ARS-UCD1.2 *Bos taurus* (GCA_002263795.2) and Brahman genomes (Ross 2019) using minimap2 v2.28 (Li 2018). The host sequences were removed using SAMtools v1.13 (Danecek *et al.* 2021), followed by alignment to the reference file with 12 specific methanogen genomes. Taxonomic classification of all non-host data was performed by Kraken2 v 2.1.2 (Wood *et al.* 2019) using the host-removed reads as input under a customised database. The NCBI RefSeq complete genomes of bacteria, fungi, archaea, and protozoa were used for database construction. The calculation of the Shannon index (alpha diversity) was performed by Vegan v2.6-8 and phyloseq v1.48.0 (McMurdie and Holmes 2013). The host sequence removal and methanogen capture efficiency, read length N50, Shannon index, and microbial relative abundance between the two adaptive protocols were compared using the regular protocol as a benchmark. The effects (*P*-values) of adaptive modes, technical replicates, and biological replicates were estimated using Fitting Linear Models.

RESULTS AND DISCUSSION

Sequencing produced 35 Gb of data in total. The filtering process removed 110 Mb (0.31%) from reads under 100 bp. The ONT depletion strategy significantly reduced the amount of host sequence data ($P < 0.01$; Figure 1A) compared to the enrichment. This indicated that the depletion protocol was a more effective alternative in sequencing host-contaminated microbiome samples, consistent with studies of other host-associated microbiomes (Ong *et al.* 2022). Depletion modes demonstrated greater N50 values compared to the enrichment method ($P < 0.001$; Figure 1B), indicating that the enrichment method rejected a substantial proportion of reads during sequencing, which could lead to the distortion of the microbial composition profile (Marquet *et al.* 2022).

The Shannon index of reads assigned to Archaea by Kracken2 was compared. The species diversity of the enrichment protocol was significantly less than the depletion mode ($P < 0.001$; Figure 1C). This indicates that the enrichment method does capture most of the 12 selected methanogen reads, a large proportion of other archaeal reads were off-target, which could be largely disadvantageous in microbiome studies where which species are present is largely unknown.

Host-removed sequences were aligned to the reference file with the 12 selected methanogen genomes. Although the enrichment function was enabled, these samples showed significantly lower percentages of methanogen bases than the depletion protocol ($P < 0.01$; Figure 1A). However, there was no notable difference in the relative abundances for the chosen archaeal genera between the two adaptive protocols in the taxonomic classification analysis ($P > 0.05$; Figure 1D). On the other hand,

the enrichment showed notably greater relative abundances of several bacterial genera, such as *Clostridium*, a trend not observed in the depletion protocol ($P < 0.001$), which indicated that the enrichment sequencing might not be a suitable protocol for comprehensive microbial profiling. The distorted bacterial composition in the enrichment may result from the false-negative rejection and high short-read proportions (Martin *et al.* 2022). In addition, these findings also indicate that the depletion strategy can preserve the original microbial distribution in the sample (Marquet *et al.* 2022), which enables the effective microbial profile characterisation with a much smaller dataset.

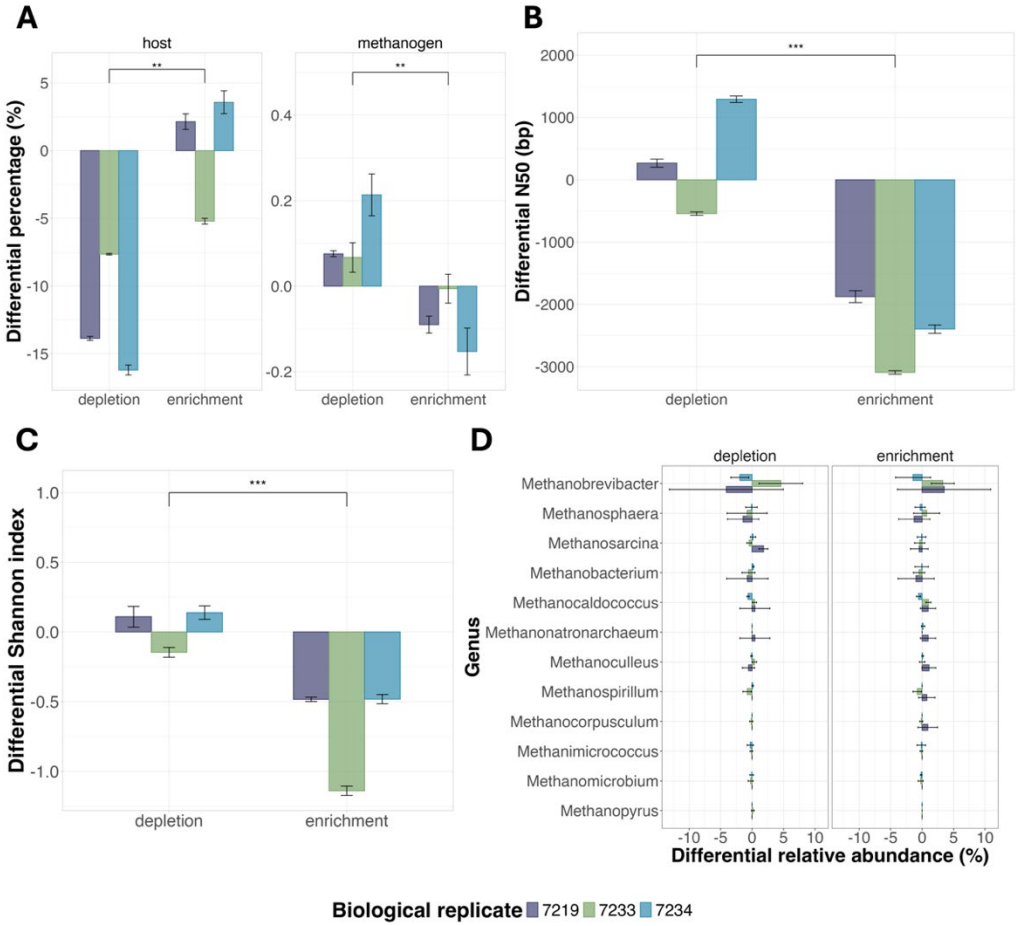

Figure 1. A) Differential proportions of bases aligned to the host genomes and selected methanogen sequences, B) differential read length N50, C) differential Shannon index (alpha diversity) of the archaeal subset, and D) differential relative abundance of selected methanogen genera between the depletion and enrichment protocols. $P < 0.05$: *; $P < 0.01$: **; $P < 0.001$: ***.

The host sequence removal and methanogen capture efficiency, read length N50, Shannon index, and microbial relative abundance between the two adaptive protocols were compared using the regular protocol as a benchmark. Technical replicates of each biological replicate were used to calculate the mean and standard deviation. All values are displayed as differences between the adaptive sequencing results, and the regular protocol results from the same biological sample averaged across the three technical replicates.

CONCLUSION

The ONT depletion protocol efficiently reduced the host-derived bases in the sequencing dataset and revealed a higher archaeal diversity compared to the enrichment strategy. The adaptive sequencing protocols did not affect the selected archaeal genus profile. These findings indicated that the ONT depletion function may reduce the high sequencing depth requirement in microbiome analysis. Therefore, ONT depletion sequencing is an efficient and cost-effective alternative to minimise host contamination in saliva microbiome data while still retaining the original microbial profile.

ACKNOWLEDGEMENTS

We appreciate the valuable support from our colleagues at the Centre for Animal Science within the Queensland Alliance for Agriculture and Food Innovation. We acknowledge the Low Emission Saliva Test for Ruminants project (LESTR), founded by Meat & Livestock Australia, grant number P.PSH.2010. We also thank the Graduate School Scholarships from the University of Queensland.

REFERENCES

Danecek P., Bonfield J.K., Liddle J., Marshall J., Ohan V., Pollard M.O., Whitwham A., Keane T., Mccarthy S.A., Davies R.M. and Li H. (2021) *Gigascience.* **10**: giab008.

De Coster W., D'hert S., Schultz D.T., Cruts M. and Van Broeckhoven C. (2018) *Bioinformatics* **34**: 2666.

Ganda E., Beck K.L., Haiminen N., Silverman J.D., Kawas B., Cronk B.D., Anderson R.R., Goodman L.B. and Wiedmann M. (2021) *Msystems.* **6**: e0061921.

Li H. (2018) *Bioinformatics* **34**: 3094.

Marquet M., Zöllkau J., Pastuschek J., Viehweger A., Schleussner E., Makarewicz O., Pletz M.W., Ehricht R. and Brandt C. (2022) *Sci Rep.* **12**: 4000.

Martin S., Heavens D., Lan Y.X., Horsfield S., Clark M.D. and Leggett R.M. (2022) *Genome Biol.* **23**: 11.

Mcmurdie P.J. and Holmes S. (2013) *Plos One* **8**: e61217.

Ong C.T., Ross E.M., Boe-Hansen G.B., Turni C., Hayes B. and Tabor A.E. (2022) *J Anim Sci.* **100**: skab344.

Ross E. (2019) Final Report P.PSH.0868, Meat and Livestock Australia Limited Locked Bag 1961 North Sydney, Australia.

Wood D.E., Lu J. and Langmead B. (2019) *Genome Biol.* **20**: 257.