TITLE: Long reads from Nanopore sequencing as a tool for animal microbiome studies

Long reads in livestock microbiota

Beatriz Delgado[1,2], Magdalena Serrano[1], Carmen González[1], Alex Bach[3,4], and Oscar González-Recio[1,2]

[1]Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, 28040 Madrid, Spain.

[2]Departamento de Producción Agraria, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain.

[3]Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain

[4]Department of Ruminant Production, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), 08140 Caldes de Montbui, Spain

Correspondence: Oscar González-Recio, Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, 28040 Madrid, Spain. Phone Number: (34) 913471486, gonzalez.oscar@inia.es

Abstract

In the era of bioinformatics and metagenomics, the study of the ruminal microbiome has gained considerable relevance in the field of animal breeding, since the composition of the rumen microbiota significantly impacts production and the environment. Illumina sequencing is considered the gold standard for the analysis of microbiomes, but it is limited by obtaining only short DNA sequences to analyze. As an alternative, Oxford Nanopore Technologies (ONT) has developed a new sequencing technique based on nanopores that can be carried out in the MinION, a portable device with a low initial cost which long DNA readings can be obtained with. The aim of this study was to compare the performance of both types of sequencing applied to samples of ruminal content using a similar pipeline. The ONT sequencing provided similar results to the Illumina sequencing, although it was able to classify a greater number of readings at the species level, possibly due to the increase in the read size. The results also suggest that, due to the size of the reads, it would be possible to obtain the same amount of information in a smaller number of hours. However, detection of archaeal and eukaryotic species is still difficult to accomplish due to their low abundance in the rumen compared to bacteria, suggesting different pipelines and strategies are needed to obtain a whole representation of the less abundant species in the rumen microbiota.

Background

The bovine rumen has been studied for years in an attempt to reveal functions and microorganisms associated with nutritional features such as feed efficiency to use them in animal breeding programs for cattle (Knapp, Laur, Vadas, Weiss, & Tricarico, 2014). Lately, methane emissions from ruminants are among the main concerns in animal husbandry, given their contribution to global warming. Microbial cultures were essential for the first descriptions of the rumen content but they are usually hard to achieve (Creevey, Kelly, Henderson, & Leahy, 2014). Thanks to the Next Generation Sequencing (NGS) techniques it is now possible to detect different microbial taxa in rumen samples avoiding culture (Seshadri et al., 2018). This has allowed analyzing metagenomic samples in an easier way and to detect non-culturable microbes that broaden the knowledge of complex microbial communities as well as allow quantifying the relative abundances of each one of them in the community.

The sequencing by synthesis method utilized by the Illumina platform is currently widely used (Goodwin et al., 2016) and accepted as a gold standard because, among other, its high sequence accuracy. The nanopore sequencing developed by Oxford Nanopore Technologies (ONT) is an attractive technology due to the availability of small and portable devices currently possible at affordable costs (Lu, Giordano, & Ning, 2016). Combined with the diverse kits offered, the ONT is a versatile alternative that can be used for multiple purposes in microbiome studies. There are some publications comparing Illumina and nanopore sequencing (Shin et al., 2016), although most of them focus on 16s rRNA gene sequencing (Cusco et al., 2017; Ma, Stachler, & Bibby, 2017). These previous studies showed a promising potential of nanopore sequencing for species detection with a high correlation between the results obtained by Illumina and ONT at the phylum and genus levels, despite the smaller basecalling accuracy from ONT compared with Illumina. However, latest ONT technology and software developments provide consensus accuracies at genome assembly larger than 99% (Wick, Judd, & Holt, 2019). Nanopore long reads are also being used to assembly whole genomes combining them with short reads for a better accuracy, and the number of microbiome studies using ONT is constantly growing. The clown anemofish, *Amphiprion ocellaris*, has been recently sequenced using hybrid assembly (Tan et al., 2018) as well as a multidrug-resistant COL1 strain (Xia et al., 2017), whose plasmids were also sequenced using short and long reads. This study also showed the ONT potential to make resistome profiles in municipal sewage, obtaining similar results as Illumina, although it was still necessary to improve the throughput of the sequencing runs. ONT has also been used to detected arbovirus in mosquitos (Batovska, Lynch, Rodoni, Sawbridge, & Cogan, 2017) obtaining comparable results to Illumina sequencing, despite the poorer quality of the reads was poorer. Other studies focus on the applicability of ONT for real-time viral pathogen detection (Greninger et al., 2015) in human blood as it shows similar results to Illumina but provides faster results, which is essential in the field of medicine. Moreover, analysis of the 16S rRNA gene is still preferred over shotgun metagenomics using the MinION and most of these studies have been conducted on mouse (Shin et al., 2016) and human gut microbiota (Leggett et al., 2017), proving consistent results for

74 individual taxonomic profiles. Other animal microbiomes have also been sequenced with the MinION, the dog
75 skin microbiota was characterized using a mock community and dog skin samples to test new 16s rRNA
76 primers, finding accurate taxonomic results at the genus level (Cusco et al., 2017).
77
78 There is a lack of publications comparing the performance of nanopore sequencing using shotgun sequencing
79 rather than 16s rRNA gene sequencing applied to non-human microbiomes such as the rumen microbiota.
80 Furthermore, to our knowledge, there are still no publications comparing ONT and Illumina to sequence
81 rumen microbiota. Assessing the ONT performance in animal microbiota is relevant as there are several
82 applications for animal husbandry. For instances, nutrition and breeding strategies are being developed to
83 modulate the microbiome in several species (ref). It may be possible to use this technology to detect favorable
84 or pathogen microbial profiles for economically and environmentally relevant traits.
85
86 Hence, the objective of this study was to evaluate the performance of ONT using the MinION (Oxford
87 Nanopore Technology, Oxford, UK) device to sequence the rumen microbiota and compare the results with an
88 Illumina MiSeq platform as benchmark in terms of reads yield, taxonomic assignment, and alpha and beta-
89 diversity.
90
91 **2. Methods**
92
93 **2.1 Sample collection**
94 Rumen content (50 µL) from 80 animals was extracted using a stomach tube connected to a mechanical
95 pumping unit and was immediately frozen and stored at -80 ºC to avoid microbial growth and to preserve the
96 sample from degradation. Samples were later thawed and ground until solid and liquid phases were
97 homogenized using a blender. DNA from 250 µL of each sample was extracted using the "DNeasy Power Soil
98 Kit" (QIAGEN, Valencia, CA, USA). DNA samples of 12 of those animals were selected for extreme
99 phenotype of feed efficiency (FE), calculated as milk production (kg/d) divided by feed consumption (kg/d), 6
100 for each type of phenotype of feed efficiency. Each sample was analyzed for quality measurement using a
101 Nanodrop ND-1000 UV/Vis spectrophotometer (Nanodrop Technologies Inc., DE, USA), recording their
102 260/280 and 260/230 ratios.
103
104 **2.2 DNA sequencing**
105 For Illumina sequencing, samples were diluted to a concentration of 5 ng/µL in a total volume of 15
106 µL/sample in a 96-well plate and were later sequenced using Illumina MiSeq in an external sequencing
107 service (FISABIO, Valencia, Spain). The quality control was performed by the same sequencing service using
108 the prinseq-lite program (Schmieder & Edwards, 2011) and the forward- and reverse-reads from the
109 sequencing were joined using the FLASH software (Magoč & Salzberg, 2011).

The MinION device was used for ONT sequencing. The Ligation Sequencing Kit (SQK-LSK109 for 1D metabarcoding) was used for the library preparation as described by the manufacturer. Reads were basecalled and demultiplexed using Guppy (community.nanoporetech.com). The resulting FASTQ files were used as input of Porechop (Wick R., 2017) for adaptor removal and barcode de-multiplexing. A total of 7,111,651 paired reads were obtained from Illumina sequencing after the QC (Table 1). After the QC for ONT reads, a total of 2,927,404 good quality reads with an average length of 1,885 bp (Table 1) were retained and the distribution of the reads among the barcodes was homogeneous.

Illumina sequence data are available from NCBI database, with bioproject number PRJNA423103. ONT sequences are available from the authors upon reasonable request.

2.3 Data analysis

The FASTQ files were aligned against the NCBI-nr protein database (Nov. 2017) using DIAMOND v0.9.22 (blastx option) setting the -F option to 15, to consider frame-shift errors in the sequences and the –rangeCulling and –top options set to 10 to scan the whole sequence for alignments with a 10% of the best local bit score (megan.informatik.uni-tuebingen.de, accessed on October 2018). The taxonomic binning of short reads from Illumina was performed using the daa2rma program from MEGAN Community Edition (CE) v6.11. with the option -a2t, to map the reads to the NCBI-taxonomy mapping file containing protein accessions (May 2017). Long reads from ONT were analyzed with specific parameters for long reads -lg (long reads) and -alg set to "longReads" (Huson et al., 2018). Relative taxonomic abundances were obtained for each samples and platform representing the number of reads assigned to each taxon. Relative abundances, alpha (Shannon and Simpson indexes) and beta diversity -using Bray-Curtis dissimilarity- were analyzed using the phyloseq (McMurdie & Holmes, 2013), vegan (Oksanen et al., 2019), and microbiome R (Leo Lahti et al., 2012-2019) packages.

2.4 Assembly

Short reads from the Illumina dataset were assembled using Megahit (Li, Liu, Luo, Sadakane, & Lam, 2015) with the default parameters using the joined forward and reverse reads. The nanopore reads were assembled using Canu (Koren et al., 2017) setting the options minReadLength = 100, minOverlapLength = 100 as the resulting reads were too short for the default parameters (less than 10,000 bp), and genomeSize = 2.5m, as described in the documentation for metagenome assemblies. The quality of the assemblies was assessed using Quast v4.0 (Mikheenko, Valin, Prjibelski, Saveliev, & Gurevich, 2016).

3. Results

3.1 Rarefaction curves

Nanopore sequencing had a low throughput in terms of number of reads when compared with Illumina due to the sequencing technology itself. On the other hand, ONT sequencing obtained longer reads as opposed to Illumina. However, rarefaction curves showed that intra-sample diversity was well represented for both technologies (Supplementary Figure 1), suggesting that a low number of reads from nanopore sequencing is enough to represent the microbial complexity in rumen samples. The 48-h sequencing runs in the MinION could be shortened, thus allowing an even faster analysis.

3.2 Alpha diversity

Alpha-diversity indexes -Shannon and Simpson- showed a high similarity for each sample in both sequencing datasets, although samples ranked slightly different between platforms (Figure 1). However, the average indexes by platform are highly similar for both the Illumina and ONT datasets -3.20 and 3.30 for the Shannon index respectively, and 0.912 and 0.918 for the Simpson index. The platform effect was not significant ($P>0.05$) when it was included in a generalized linear model with Shannon or Simpson diversity indices as dependent variable, and parity and platform as explanatory effects.

3.3 Beta diversity

Nanopore and Illumina sequencing also delivered similar results for microbial composition and abundance at all taxonomic levels, although the nanopore dataset detected a slightly greater diversity (Table 1). Nanopore sequencing seems to detect a greater ecological diversity in the samples at lower taxonomic levels (e.g. species and genera). Longer reads allow DIAMOND and MEGAN assigning more sequences to the reference database as they rely on sequence similarity, and it is more likely to match a read to a species when using longer reads. Both platforms display similar results at upper taxonomic levels -phylum, family and genus- as shown by the similar distribution found in the ordination plots, despite the differences in the number of reads. Inter-sample relative abundances are similar for both sequencing techniques although there were some differences between MinION and Illumina datasets. These differences are greater at the species level and seem to affect sample distribution between both sequencing platforms, meaning the relative abundances of the same sample differ for the different platforms.

3.4 Relative abundances

Bacteria, Archaea and Eukarya kingdoms were detected in the expected ratios, with Bacteria being the most abundant taxonomic kingdom. Illumina reads were grouped in 125 species, 78 genera and 17 phyla (Table 1), whereas a larger number of groups were detected from ONT reads: 146 species, 86 genera and 21 phyla.

Relative abundances for common taxa were similar, although the MinION detected a greater abundance of the phylum Firmicutes (Figure 3) while maintaining the abundances of other taxa such as Bacteroidetes, which also affects the Bacteroidetes/Firmicutes ratio of each sample (Table 2, Supplementary Figure 2), creating high discrepancies in samples such as number 2631 and 3404. Moreover, a higher diversity was detected with longer reads, since 95 Firmicutes species and 50 Bacteroidetes species could be classified with ONT, whereas when using Illumina reads only yield 72 Firmicutes and 39 Bacteroidetes species respectively could be classified. ONT classified 16% of reads to species level, 2 percentual points larger than Illumina. For comparison, the Bacteroidetes/Firmicutes ratio was analyzed from 16S Illumina sequences (V3-V4 regions). Clear differences were observed in this ratio between Illumina shotgun and Illumina 16S amplicon sequencing (Supplementary Figure 2). The Bacteroidetes/Firmicutes ratio from ONT sequencing overlapped with both Illumina strategies.

3.5 Assembly

Assembly of long and short reads was performed to test the ability of nanopore reads to assembly longer contigs. The assembly of the MinION reads provided a lower number of contigs (281 contigs) compared to the Illumina assembly (7856 contigs), while also having a higher N50 (see Table 3). The total length of the assembly is different -Illumina almost doubles Nanopore's- which complicates the comparison in terms in the information each assembly provides, but this can be mainly caused by the difference in the number of reads provided by each platform. However, the higher N50 and the low number of contigs from ONT indicate that long reads outperform short reads in metagenomic assemblies, also allowing the assembly of ultra-long contigs -49,178 bp was the longest contigs in the nanopore assembly while Illumina's longest contig was only 11,173 bp. Sequencing depth with Illumina MiSeq does not seem to be large enough to obtain an assembly with enough quality.

4. Discussion

Relative abundance results for common taxa to both platforms were consistent with the previous analyses, finding a high similarity between the two platforms, although the MinION detected a greater abundance of the phylum Firmicutes (Figure 3) while maintaining the abundances of other taxa. Members of the phylum Bacteroidetes have been described as highly efficient polysaccharide degraders due to large genomic regions specialized in this processes (Seshadri et al., 2018) whereas members of Firmicutes are considered "nutritionally fastidious" as they have lost most of their degradative enzymes and feed on fermentation products produced by other microbes. Moreover, the Bacteroidetes/Firmicutes (B/F) ratio has been associated with obesity in some human and mice gut microbiota studies (Koliada et al., 2017), emphasizing that high ratios are related to a higher body mass index (BMI). Obtaining accurate B/F ratios could be key to analyze the relationship between the rumen microbiota composition and relevant phenotypic traits and more efforts are needed to sequence these microbes using nanopore technologies. Other studies in dairy cattle showed that this

ratio was strongly associated with milk-fat yield (Jami, White, & Mizrahi, 2014) and suggest that these two phyla could also be related to different feed efficiency parameters such as the dry matter intake (DMI) and residual feed intake (RFI). *Prevotella* spp. -which belongs to Bacteroidetes- has been associated to feed efficiency (Jewell, McCormick, Odt, Weimer, & Suen, 2015; Bach et al., 2019), although most of these results were not consistent due to the low taxonomic resolution of 16S rRNA gene sequencing. Species among this genus are involved in central rumen processes such as pectin and xyloglucan degradation and acetate production (Seshadri et al., 2018), which have a high impact on the nutrition of ruminants. Some previous studies showed that *Prevotella bryantii* -also detected by the two platforms- has some relationship with feed efficiency (Elolimy, Arroyo, Batistel, Iakiviak, & Loor, 2018). The similarities observed for the principal taxonomic groups imply that ONT reads could be used in association analyses to correlate the microbiome composition to phenotypic traits such as feed efficiency. A correlation analysis between the rumen microbiome and feed efficiency has already been performed using Illumina reads, where significant correlations between a high feed efficiency were found when the microbiome composition had a high abundance of Bacteroidetes and a low abundance of archaea (Delgado et al., 2019), which the current study found in almost the same proportion in both datasets. Moreover, the higher number of species classified by long reads indicates that it is likely that nanopore reads are also suitable for other applications such as pathogen detection in cattle since its higher species resolution.

The overall diversity of methanogens assigned by the ONT reads was low, detecting only two species belonging to the genus *Methanobrevibacter*. The species *M. millerae* was the only common species between the two platforms, while the Illumina reads also included other important species such as *M. ruminantium* and *M. olleyae*. *Methanobrevibacter spp*. In addition, relevant ciliates appear in similar proportions in both datasets, such as *Entodinium caudatum, Tetrahymena thermophila, Paramecium tetraurelia* and *Oxytricha trifallax*, although the presence of these microorganisms was not consistent since some samples had no eukaryotic representatives, which happened both in the ONT and the Illumina dataset. The low detection of important archaeal species such as *M. ruminantium* and *M. olleyae* are one of the most relevant differences between the two platforms. *Methanobrevibacter spp*. is usually the main methanogens in rumen samples and several studies are consistent in the presence of *M. millerae* -which was found in both datasets- and *M. ruminantium* (Chaucheyras-Durand & Ossa, 2014). Moreover, the rumen methanogen *M. millerae* M9 has been shown to have more copies of methanogenesis-related genes (Kelly et al., 2016) and some members of the class Thermoplasmata -also detected by the Illumina reads- have been proposed as a target for methane reduction (Poulsen et al., 2013) as they are methylotrophic and do not require hydrogen to synthesize methane. The ciliates detected by both platforms -*Entodinium caudatum, Tetrahymena thermophila, Paramecium tetraurelia* and *Oxytricha trifallax*-,are also commonly found in the rumen microbiome (Newbold, De la Fuente, Belanche, Ramos-Morales, & McEwan, 2015). Ciliates and methanogens are usually symbionts and are involved in methane production, although only methanogens produce methane (Newbold et al., 2015). Members of Ciliophora have more hydrogenosomes than other protozoa and are believed to have

more endosymbiotic methanogens than other protozoa and to produce a greater impact on methanogenesis (Newbold et al., 2015). However, presence of eukaryotic microorganisms was not detected in some samples, which can be due to the low abundance of this species, compared with bacteria in the rumen microbiota or due to sensitivity issues in the pipeline. These results suggest that the taxonomic classification is comparable to that of Illumina at the phylum, family and genus levels, although long reads can be useful for species detection (e.g. pathogen detection). More sensitive procedures, such as amplicon or operon sequencing, may be necessary for a better detection of low abundant microorganisms. Increasing the sequencing depth could also be useful to detect a higher number of non-abundant species (Sims, Sudbery, Ilott, Heger, & Ponting, 2014) such as archaea and eukaryotes in rumen samples. Better extraction methods that avoid DNA fragmentation may improve the number of long-DNA molecules that are recovered from the samples, which also allows a better nanopore sequencing.

Assembly of metagenomic data is the best strategy to recover genomes from the data, while non-assembly strategies are usually better to analyze microbiome profiles. Moreover, assemblies are useful for functional profiling and for connecting relevant metabolic pathways to their corresponding microbes and for discovering and characterizing important species. Although this study suggests that nanopore sequencing outperforms short-read assemblies for more accuracy in the final assemblies, a combination of both types of reads is usually the best way to obtain high quality genomes (Tan et al., 2018), as Illumina's short reads are usually highly accurate. However, nanopore reads are expected to reach the same accuracies as new bioinformatic tools are being developed so high-quality assemblies could be obtained using only nanopore reads, although the sequencing depth plays an important role in the quality of the assemblies, and it should be increased to obtain better results. The average size of the reads obtained by the MinION must be longer for better assemblies, as they would cover larger parts of the genomes. For that, the protocols for DNA extraction should be improved to avoid shearing the microbial DNA.

Classical approaches for modulating the microbiome include management practices and food additives, genetic selection is the principal strategy to obtain animals with improved phenotypic traits (Tapio, Snelling, Strozzi, & Wallace, 2017). It has been suggested that the microbial composition in ruminants could be used as a predictor of complex traits (Gonzalez-Recio, Zubiria, García-Rodríguez, Hurtado, & Atxaerandio, 2018), as there are signs of host genetic control over the microbiome composition in cows, which could lead to animal breeding programs that select animals with a favorable microbiome for a high feed efficiency phenotype (Sasson et al., 2017). Thus, selecting animals with less methane production could lead to a decrease in the overall impact of livestock in our environment.

Conclusions

Illumina sequencing is the most used sequencing technology for shotgun metagenomics as it has a high sequencing depth and accuracy, but the short length of the reads makes it difficult to assign specific taxa or

295 genes without a previous assembly. Nanopore sequencing provides long reads that could cover larger genome
296 regions allowing more accurate assemblies. However, long reads involve greater computational efforts. Here
297 we used DIAMOND and MEGAN tools which can work with long reads. Our analyses suggest that ONT
298 sequencing provide an interesting alternative for microbiome profiling in rumen samples following this
299 protocol. The microbiome compositions determined using the two technologies are similar, especially at the
300 phylum and family levels, although a larger number of taxonomic groups is detected when using longer reads.
301 Both platforms provide similar beta-diversity between sequenced samples. The ONT sequencing provides an
302 easier approach to microbiome profiling for almost all kind of laboratories due to its small size, flexibility,
303 and relative low cost. The ONT could even be used in the field to analyze microbiome profiles in real time
304 and provide quick results that may be used for diagnostic or phenotyping purposes.
305
306
307 **Bibliography**
308
309 Bach, A., Lopez-Garcia, A., Gonzalez-Recio, O., Elcoso, G., Fàbregas, F., Chaucheyras-Durand, F., & Castex
310 M. (2019). Changes in the rumen and colon microbiota and effects of live yeast dietary supplementation
311 during the transition from the dry period to lactation of dairy cows. *Journal of Dairy Science, 102(7):*
312 *6180-6198*. https://doi.org/10.3168/jds.2018-16105
313 Batovska, J., Lynch, S. E., Rodoni, B. C., Sawbridge, T. I., & Cogan, N. O. (2017). Metagenomic arbovirus
314 detection using MinION nanopore sequencing. *Journal of Virological Methods*, *249*(September), 79–84.
315 https://doi.org/10.1016/j.jviromet.2017.08.019
316 Chaucheyras-Durand, F., & Ossa, F. (2014). The rumen microbiome : abundance , diversity , and new
317 investigative tools. *The Professional Animal Scientist*, *30*, 1–12.
318 https://doi.org/http://dx.doi.org/10.15232/S1080-7446(15)30076-0
319 Creevey, C. J., Kelly, W. J., Henderson, G., & Leahy, S. C. (2014). Determining the culturability of the rumen
320 bacterial microbiome. *Microbial Biotechnology*, *7*(5), 467–479. https://doi.org/10.1111/1751-
321 7915.12141
322 Cusco, A., Vines, J., D'Andreano, S., Riva, F., Casellas, J., Sanchez, A., & Francino, O. (2017). Using
323 MinION to characterize dog skin microbiota through full-length 16S rRNA gene sequencing approach.
324 *BioRxiv*, 167015. https://doi.org/10.1101/167015
325 Delgado, B., Bach, A., Guasch, I., González, C., Elcoso, G., Pryce, J. E., & Gonzalez-Recio, O. (2019).
326 Whole rumen metagenome sequencing allows classifying and predicting feed efficiency and intake
327 levels in cattle. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-018-36673-w
328 Elolimy, A. A., Arroyo, J. M., Batistel, F., Iakiviak, M. A., & Loor, J. J. (n.d.). *Association of residual feed*
329 *intake with abundance of ruminal bacteria and biopolymer hydrolyzing enzyme activities during the*
330 *peripartal period and early lactation in Holstein dairy cows*. https://doi.org/10.1186/s40104-018-0258-9
331 Gonzalez-Recio, O., Zubiria, I., García-Rodríguez, A., Hurtado, A., & Atxaerandio, R. (2018). Short

communication: Signs of host genetic regulation in the microbiome composition in 2 dairy breeds: Holstein and Brown Swiss. *Journal of Dairy Science*, *101*(3), 2285–2292. https://doi.org/10.3168/jds.2017-13179

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), 333–351. https://doi.org/10.1038/nrg.2016.49

Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., … Chiu, C. Y. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Medicine*. https://doi.org/10.1186/s13073-015-0220-9

Huson, D. H., Albrecht, B., Bağci, C., Bessarab, I., Górska, A., Jolic, D., & Williams, R. B. H. (2018). MEGAN-LR: New algorithms allow accurate binning and easy interactive exploration of metagenomic long reads and contigs. *Biology Direct*. https://doi.org/10.1186/s13062-018-0208-7

Jami, E., White, B. A., & Mizrahi, I. (2014). Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. *PLoS ONE*, *9*(1). https://doi.org/10.1371/journal.pone.0085423

Jewell, K. A., McCormick, C. A., Odt, C. L., Weimer, P. J., & Suen, G. (2015). Ruminal bacterial community composition in dairy cows is dynamic over the course of two lactations and correlates with feed efficiency. *Applied and Environmental Microbiology*. https://doi.org/10.1128/AEM.00720-15

Kelly, W. J., Pacheco, D. M., Li, D., Attwood, G. T., Altermann, E., & Leahy, S. C. (2016). The complete genome sequence of the rumen methanogen Methanobrevibacter millerae SM9. *Standards in Genomic Sciences*. https://doi.org/10.1186/s40793-016-0171-9

Knapp, J. R., Laur, G. L., Vadas, P. A., Weiss, W. P., & Tricarico, J. M. (2014). Invited review: Enteric methane in dairy cattle production: Quantifying the opportunities and impact of reducing emissions. *Journal of Dairy Science*. https://doi.org/10.3168/jds.2013-7234

Koliada, A., Syzenko, G., Moseiko, V., Budovska, L., Puchkov, K., Perederiy, V., … Vaiserman, A. (2017). Association between body mass index and Firmicutes/Bacteroidetes ratio in an adult Ukrainian population. *BMC Microbiology*. https://doi.org/10.1186/s12866-017-1027-1

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

Leggett, R. M., Alcon-Giner, C., Heavens, D., Caim, S., Brook, T. C., Kujawska, M., … Clark, M. D. (2017). Rapid MinION metagenomic profiling of the preterm infant gut microbiota to aid in pathogen diagnostics. *Bioarxiv*, 180406. https://doi.org/10.1101/180406

Leo Lahti and Sudarshan Shetty. (n.d.). *microbiome R package*. Retrieved from http://microbiome.github.io

Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics and Bioinformatics*. https://doi.org/10.1016/j.gpb.2016.05.004

Ma, X., Stachler, E., & Bibby, K. (2017). Evaluation of Oxford Nanopore MinION Sequencing for 16S rRNA Microbiome Characterization. *BioRxiv*, 1–19. https://doi.org/10.1101/099960

Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, *27*(21), 2957–2963. Retrieved from http://dx.doi.org/10.1093/bioinformatics/btr507

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, *8*(4). https://doi.org/10.1371/journal.pone.0061217

Mikheenko, A., Valin, G., Prjibelski, A., Saveliev, V., & Gurevich, A. (2016). Icarus: Visualizer for de novo assembly evaluation. *Bioinformatics*, *32*(21), 3321–3323. https://doi.org/10.1093/bioinformatics/btw379

Newbold, C. J., De la Fuente, G., Belanche, A., Ramos-Morales, E., & McEwan, N. R. (2015). The role of ciliate protozoa in the rumen. *Frontiers in Microbiology*. https://doi.org/10.3389/fmicb.2015.01313

Oksanen, A. J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., … Szoecs, E. (2019). *Package ' vegan .'*

Poulsen, M., Schwab, C., Borg Jensen, B., Engberg, R. M., Spang, A., Canibe, N., … Urich, T. (2013). Methylotrophic methanogenic Thermoplasmata implicated in reduced methane emissions from bovine rumen. *Nature Communications*. https://doi.org/10.1038/ncomms2432

Sasson, G., Kruger Ben-Shabat, S., Seroussi, E., Doron-Faigenboim, A., Shterzer, N., Yaacoby, S., … Mizrahi, I. (2017). Heritable Bovine Rumen Bacteria Are Phylogenetically Related and Correlated with the Cow's Capacity To Harvest Energy from Its Feed. *MBio*, *8*(4). Retrieved from http://mbio.asm.org/content/8/4/e00703-17.abstract

Schmieder, R., & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, *27*(6), 863–864. Retrieved from http://dx.doi.org/10.1093/bioinformatics/btr026

Seshadri, R., Leahy, S. C., Attwood, G. T., Teh, K. H., Lambie, S. C., Cookson, A. L., … Cerón Cucchi, M. (2018). Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nature Biotechnology*. https://doi.org/10.1038/nbt.4110

Shin, J., Lee, S., Go, M. J., Lee, S. Y., Kim, S. C., Lee, C. H., & Cho, B. K. (2016). Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing. *Scientific Reports*. https://doi.org/10.1038/srep29681

Sims, D., Sudbery, I., Ilott, N. E., Heger, A., & Ponting, C. P. (2014). Sequencing depth and coverage: Key considerations in genomic analyses. *Nature Reviews Genetics*, *15*(2), 121–132. https://doi.org/10.1038/nrg3642

Tan, M. H., Austin, C. M., Hammer, M. P., Lee, Y. P., Croft, L. J., & Gan, H. M. (2018). Finding Nemo: Hybrid assembly with Oxford Nanopore and Illumina reads greatly improves the clownfish (Amphiprion ocellaris) genome assembly. *GigaScience*, *7*(3), 1–6. https://doi.org/10.1093/gigascience/gix137

Tapio, I., Snelling, T. J., Strozzi, F., & Wallace, R. J. (2017). The ruminal microbiome associated with methane emissions from ruminant livestock. *Journal of Animal Science and Biotechnology*, *8*(1), 1–11. https://doi.org/10.1186/s40104-017-0141-0

406 Wick R. (n.d.). *Porechop*. Retrieved from https://github.com/rrwick/Porechop
407 Wick, R. R., Judd, L. M., & Holt, K. E. (2019). Performance of neural network basecalling tools for Oxford
408 Nanopore sequencing. *BioRxiv*, 543439. https://doi.org/10.1101/543439
409 Xia, Y., Li, A. D., Deng, Y., Jiang, X. T., Li, L. G., & Zhang, T. (2017). MinION Nanopore sequencing
410 enables correlation between resistome phenotype and genotype of coliform bacteria in municipal
411 sewage. *Frontiers in Microbiology*, *8*(OCT), 1–13. https://doi.org/10.3389/fmicb.2017.02105
412
413

414

415 **Table 1. Summary of Illumina and Nanopore sequencing performance and analysis results.**

Sample	Illumina reads	Nanopore reads
1147	763,635	242,794
1149	614,876	268,196
678	580,763	221,340
17246	519,271	229,538
2631	836,541	245,231
152	425,134	272,455
8172	501,871	204,676
8627	784,286	273,294
9524	575,278	307,658
8815	508,088	269,539
3404	534,667	147,489
2954	467,241	245,194
TOTAL reads	7,111,651	2,927,404
Total yield	2.13 Gb	5.48 Gb
Species	125	146
Genera	78	86
Phyla	17	21
Read length (pb)	2 x 300	≈1,885

416

417

418 **Table 2. Ratio Bacteroidetes/Firmicutes in both datasets**

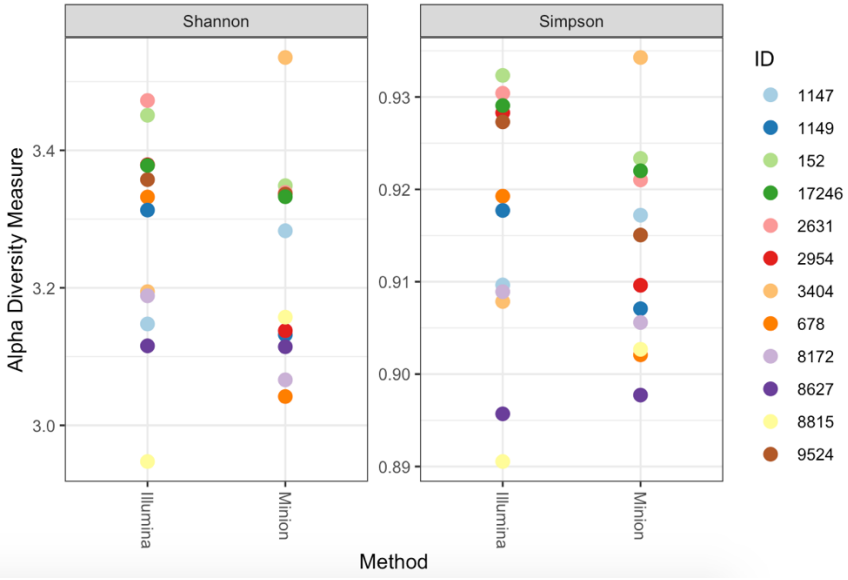
Sample	MinION ratio	Illumina ratio
2631	0.34476036	2.34547541
2954	1.77699134	2.3289202
3404	0.74890076	2.16387852
678	2.20379843	3.27502731
8172	2.38623769	2.91414699
8627	0.85845545	1.74129358

8815	1.30866685	2.37871471
9524	1.68766111	2.39366088
1147	1.48723686	1.89157119
1149	2.74632061	2.41265928
152	2.77646755	3.15533655
17246	1.97668826	1.13413476

Table 3. Assembly statistics

	Illumina assembly	MinION assembly
Length (bp)	5,296,436	2,550,082
GC (%)	47.7	46.89
N's	0	0
N50	641	11,549
Largest Contig (bp)	11,173	49,178
Contigs	7,856	281

425 **Figure 1. Comparison of alpha diversity indexes.** Shannon (left) and Simpson (right) indexes were calculated for each
426 sample and sequencing technology to compare the overall taxonomic diversity.



427

428

429

430

431

Figure 1. Comparison of relative abundances of common taxa at four different taxonomic levels using Illumina and MinION sequencers. *Relative abundances are similar for common taxa at all levels, although the phylum Firmicutes has a higher relative abundance in the MinION dataset which counts for most of the observed differences in the four graphics.*

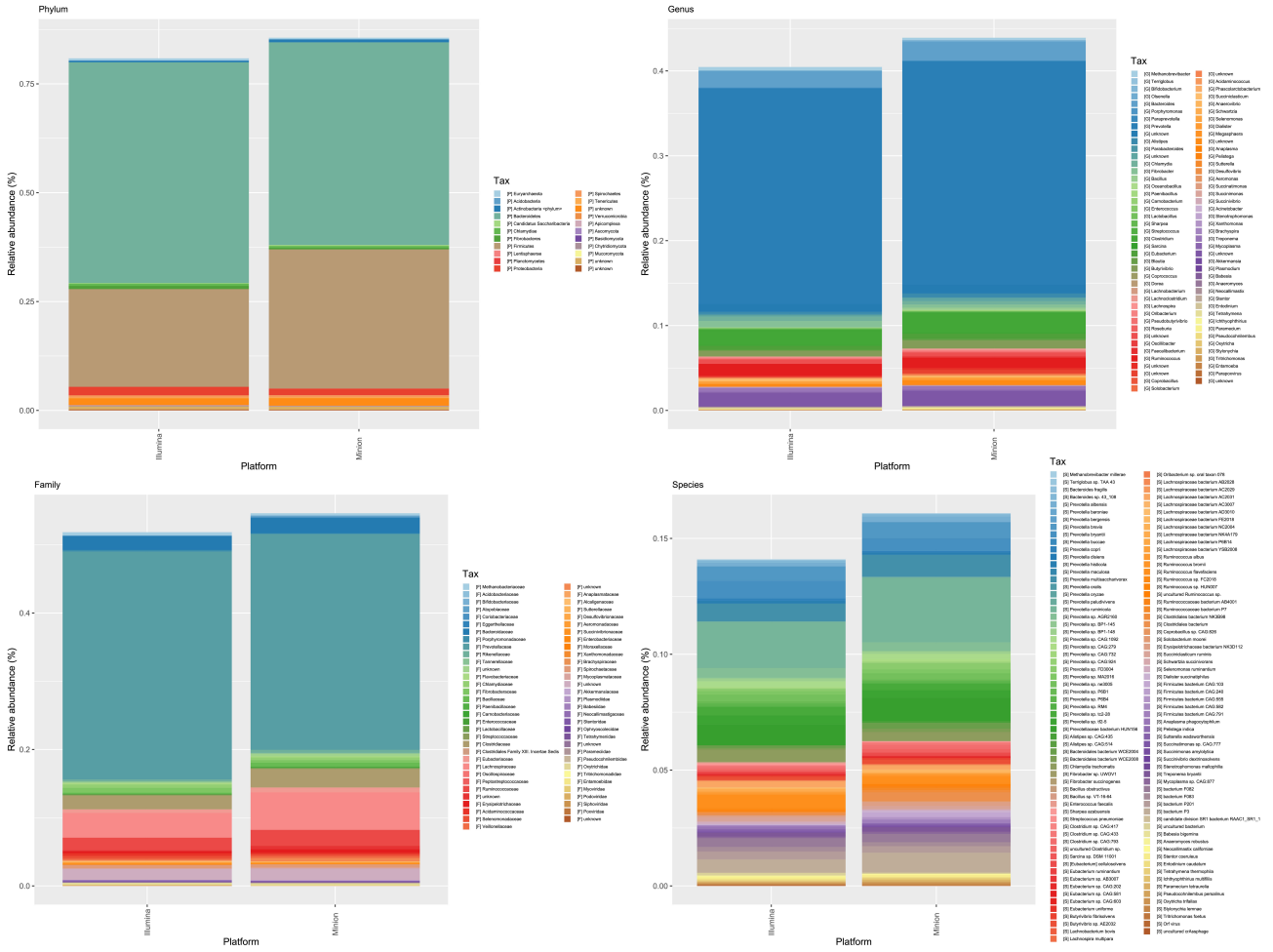
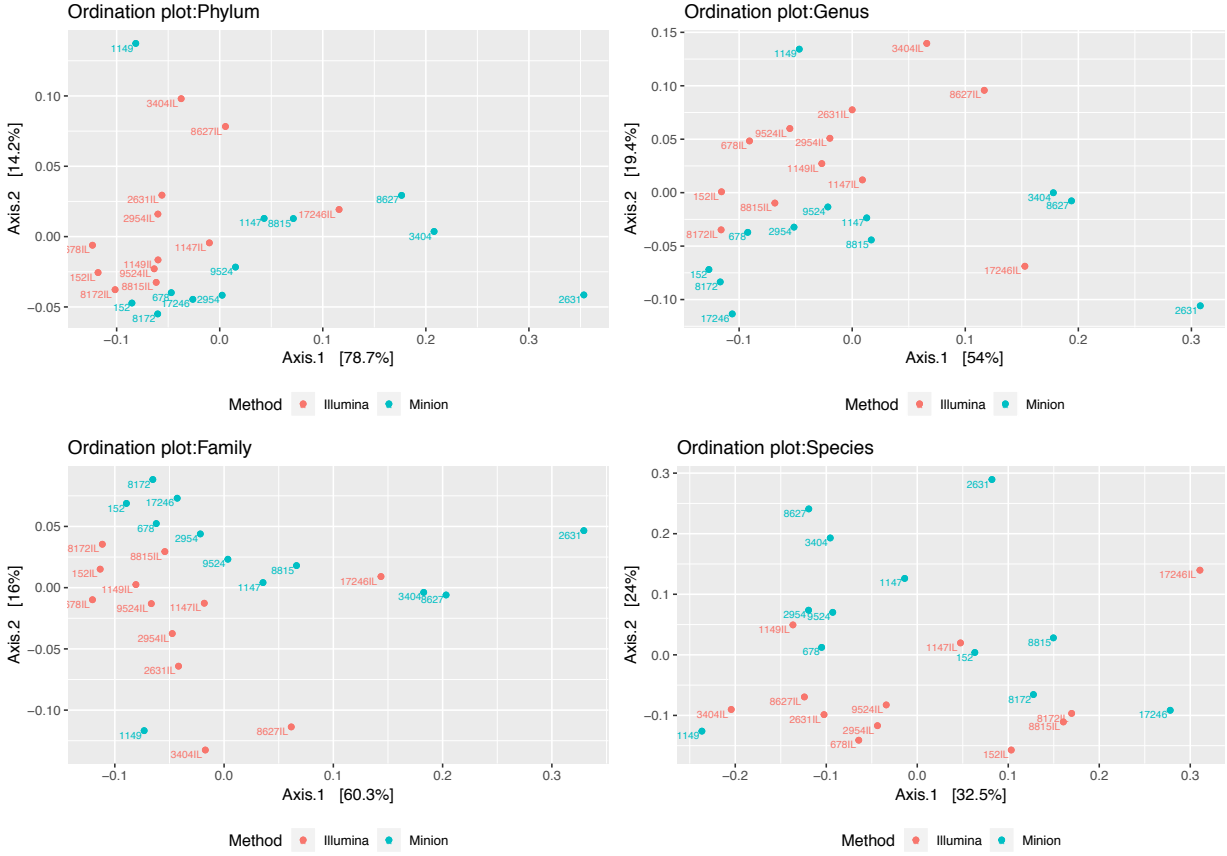


Figure 3. Ordination plots corresponding to four different taxonomic levels. Ordination analysis was performed using data corresponding to relative abundances for both sequencing techniques at different taxonomic levels: phylum, family, genus and species. For the phylum, family and genus levels two different clusters can be observed corresponding to the different sequencing techniques, although sample distribution is similar in both types of sequencing. At the species level no obvious clusters can be observed.

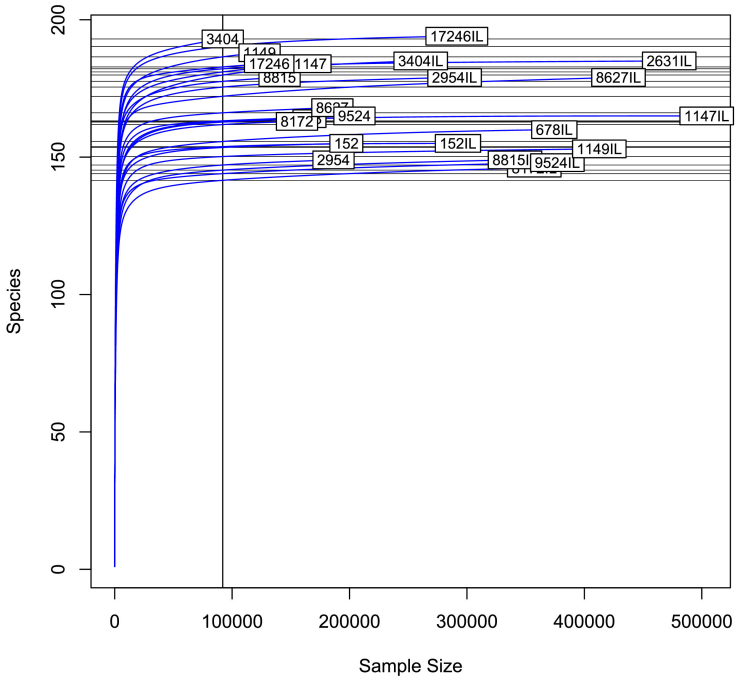


446

447 **Supplementary Figures**

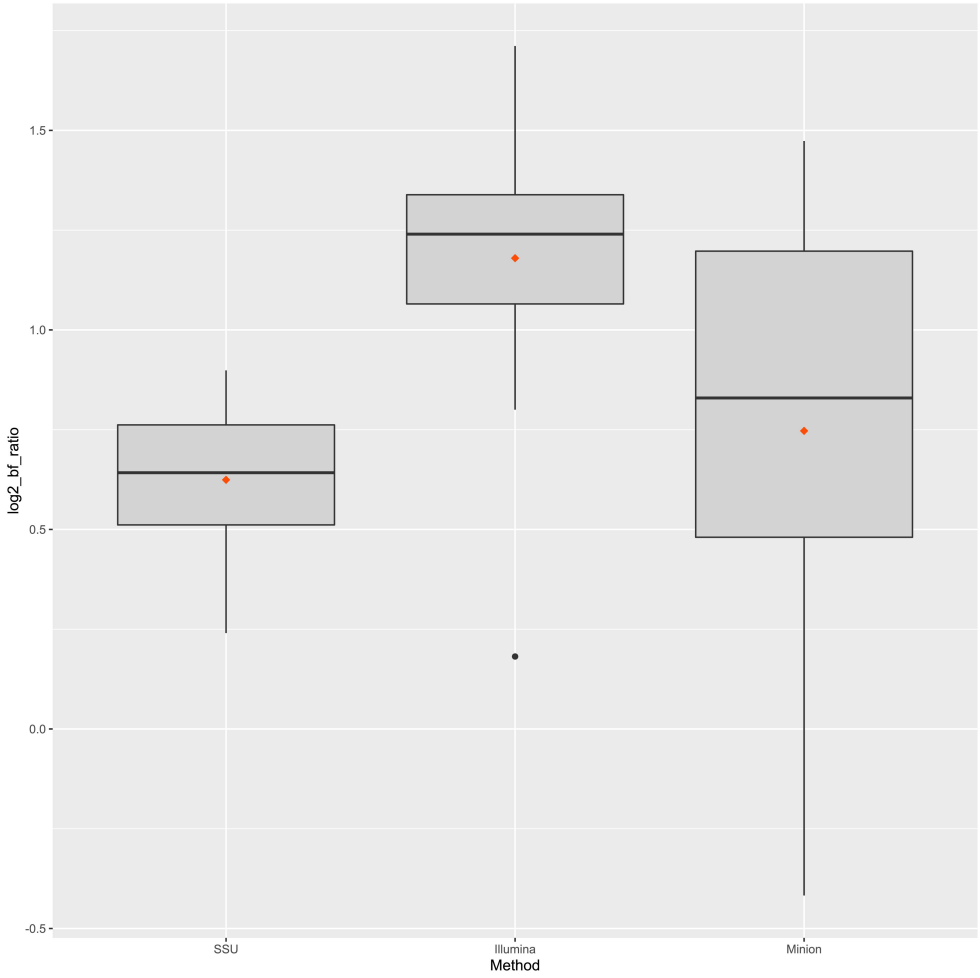
448

449 Supplementary figure 1. Rarefaction curves of the 12 samples for both sequencing techniques.



450
451
452

453 Supplementary figure 2. Average ratio Bacteroidetes/Firmicutes per platform.



454