

 Open access • Posted Content • DOI:10.1101/2019.12.21.886028

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

— [Source link](#) 

Beatriz Delgado, Magdalena Serrano, Carmen G. Gonzalez, Alex Bach ...+1 more authors

Institutions: Technical University of Madrid, Catalan Institution for Research and Advanced Studies

Published on: 23 Dec 2019 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Nanopore sequencing, Illumina dye sequencing, Metagenomics and DNA sequencing

Related papers:

- [Implications of error-prone long-read whole-genome shotgun sequencing on characterizing reference microbiomes](#)
- [Abundance tracking by long-read nanopore sequencing of complex microbial communities in samples from 20 different biogas/wastewater plants](#)
- [Ultra-accurate Microbial Amplicon Sequencing with Synthetic Long Reads](#)
- [Ultra-accurate Microbial Amplicon Sequencing Directly from Complex Samples with Synthetic Long Reads](#)
- [Accurate Microbiome Sequencing with Synthetic Long Read Sequencing](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/long-reads-from-nanopore-sequencing-as-a-tool-for-animal-3fdw7vm55z>

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

2
3 Long reads in livestock microbiota
4
5

6
7 **Beatriz Delgado^{1,2}, Magdalena Serrano¹, Carmen González¹, Alex Bach^{3,4}, and Oscar González-Recio^{1,2}**

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

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

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

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

13
14 Correspondence: Oscar González-Recio, Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y
15 Alimentaria, 28040 Madrid, Spain. Phone Number: (34) 913471486, gonzalez.oscar@inia.es
16
17

18 **Abstract**

19
20 In the era of bioinformatics and metagenomics, the study of the ruminal microbiome has gained considerable
21 relevance in the field of animal breeding, since the composition of the rumen microbiota significantly impacts
22 production and the environment. Illumina sequencing is considered the gold standard for the analysis of
23 microbiomes, but it is limited by obtaining only short DNA sequences to analyze. As an alternative, Oxford
24 Nanopore Technologies (ONT) has developed a new sequencing technique based on nanopores that can be
25 carried out in the MinION, a portable device with a low initial cost which long DNA readings can be obtained
26 with. The aim of this study was to compare the performance of both types of sequencing applied to samples of
27 ruminal content using a similar pipeline. The ONT sequencing provided similar results to the Illumina
28 sequencing, although it was able to classify a greater number of readings at the species level, possibly due to
29 the increase in the read size. The results also suggest that, due to the size of the reads, it would be possible to
30 obtain the same amount of information in a smaller number of hours. However, detection of archaeal and
31 eukaryotic species is still difficult to accomplish due to their low abundance in the rumen compared to
32 bacteria, suggesting different pipelines and strategies are needed to obtain a whole representation of the less
33 abundant species in the rumen microbiota.
34
35
36

37 **Background**

38

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

49

50 The sequencing by synthesis method utilized by the Illumina platform is currently widely used (Goodwin et
51 al., 2016) and accepted as a gold standard because, among other, its high sequence accuracy. The nanopore
52 sequencing developed by Oxford Nanopore Technologies (ONT) is an attractive technology due to the
53 availability of small and portable devices currently possible at affordable costs (Lu, Giordano, & Ning, 2016).
54 Combined with the diverse kits offered, the ONT is a versatile alternative that can be used for multiple
55 purposes in microbiome studies. There are some publications comparing Illumina and nanopore sequencing
56 (Shin et al., 2016), although most of them focus on 16s rRNA gene sequencing (Cusco et al., 2017; Ma,
57 Stachler, & Bibby, 2017). These previous studies showed a promising potential of nanopore sequencing for
58 species detection with a high correlation between the results obtained by Illumina and ONT at the phylum and
59 genus levels, despite the smaller basecalling accuracy from ONT compared with Illumina. However, latest
60 ONT technology and software developments provide consensus accuracies at genome assembly larger than
61 99% (Wick, Judd, & Holt, 2019). Nanopore long reads are also being used to assembly whole genomes
62 combining them with short reads for a better accuracy, and the number of microbiome studies using ONT is
63 constantly growing. The clown anemofish, *Amphiprion ocellaris*, has been recently sequenced using hybrid
64 assembly (Tan et al., 2018) as well as a multidrug-resistant COL1 strain (Xia et al., 2017), whose plasmids
65 were also sequenced using short and long reads. This study also showed the ONT potential to make resistome
66 profiles in municipal sewage, obtaining similar results as Illumina, although it was still necessary to improve
67 the throughput of the sequencing runs. ONT has also been used to detected arbovirus in mosquitos (Batovska,
68 Lynch, Rodoni, Sawbridge, & Cogan, 2017) obtaining comparable results to Illumina sequencing, despite the
69 poorer quality of the reads was poorer. Other studies focus on the applicability of ONT for real-time viral
70 pathogen detection (Greninger et al., 2015) in human blood as it shows similar results to Illumina but provides
71 faster results, which is essential in the field of medicine. Moreover, analysis of the 16S rRNA gene is still
72 preferred over shotgun metagenomics using the MinION and most of these studies have been conducted on
73 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 μ L) from 80 animals was extracted using a stomach tube connected to a mechanical
95 pumping unit and was immediately frozen and stored at -80 °C 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 μ L 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/ μ L in a total volume of 15
106 μ L/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).

110 The MinION device was used for ONT sequencing. The Ligation Sequencing Kit (SQK-LSK109 for 1D
111 metabarcoding) was used for the library preparation as described by the manufacturer. Reads were basecalled
112 and demultiplexed using Guppy (community.nanoporetech.com). The resulting FASTQ files were used as
113 input of Porechop (Wick R., 2017) for adaptor removal and barcode de-multiplexing. A total of 7,111,651
114 paired reads were obtained from Illumina sequencing after the QC (Table 1). After the QC for ONT reads, a
115 total of 2,927,404 good quality reads with an average length of 1,885 bp (Table 1) were retained and the
116 distribution of the reads among the barcodes was homogeneous.
117 Illumina sequence data are available from NCBI database, with bioproject number PRJNA423103. ONT
118 sequences are available from the authors upon reasonable request.

119

120

121 **2.3 Data analysis**

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

134

135 **2.4 Assembly**

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

142

143

144

145

146

147 **3. Results**

148

149 **3.1 Rarefaction curves**

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

156

157 **3.2 Alpha diversity**

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

164

165

166 **3.3 Beta diversity**

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

178

179

180 **3.4 Relative abundances**

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

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

195

196 **3.5 Assembly**

197 Assembly of long and short reads was performed to test the ability of nanopore reads to assembly longer
198 contigs. The assembly of the MinION reads provided a lower number of contigs (281 contigs) compared to
199 the Illumina assembly (7856 contigs), while also having a higher N50 (see Table 3). The total length of the
200 assembly is different -Illumina almost doubles Nanopore's- which complicates the comparison in terms in the
201 information each assembly provides, but this can be mainly caused by the difference in the number of reads
202 provided by each platform. However, the higher N50 and the low number of contigs from ONT indicate that
203 long reads outperform short reads in metagenomic assemblies, also allowing the assembly of ultra-long
204 contigs -49,178 bp was the longest contigs in the nanopore assembly while Illumina's longest contig was only
205 11,173 bp. Sequencing depth with Illumina MiSeq does not seem to be large enough to obtain an assembly
206 with enough quality.

207

208 **4. Discussion**

209

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

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

238

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

258 more endosymbiotic methanogens than other protozoa and to produce a greater impact on methanogenesis
259 (Newbold et al., 2015). However, presence of eukaryotic microorganisms was not detected in some samples,
260 which can be due to the low abundance of this species, compared with bacteria in the rumen microbiota or due
261 to sensitivity issues in the pipeline. These results suggest that the taxonomic classification is comparable to
262 that of Illumina at the phylum, family and genus levels, although long reads can be useful for species
263 detection (e.g. pathogen detection). More sensitive procedures, such as amplicon or operon sequencing, may
264 be necessary for a better detection of low abundant microorganisms. Increasing the sequencing depth could
265 also be useful to detect a higher number of non-abundant species (Sims, Sudbery, Ilott, Heger, & Ponting,
266 2014) such as archaea and eukaryotes in rumen samples. Better extraction methods that avoid DNA
267 fragmentation may improve the number of long-DNA molecules that are recovered from the samples, which
268 also allows a better nanopore sequencing.

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

281
282 Classical approaches for modulating the microbiome include management practices and food additives,
283 genetic selection is the principal strategy to obtain animals with improved phenotypic traits (Tapio, Snelling,
284 Strozzi, & Wallace, 2017). It has been suggested that the microbial composition in ruminants could be used as
285 a predictor of complex traits (Gonzalez-Recio, Zubiria, García-Rodríguez, Hurtado, & Atxaerandio, 2018), as
286 there are signs of host genetic control over the microbiome composition in cows, which could lead to animal
287 breeding programs that select animals with a favorable microbiome for a high feed efficiency phenotype
288 (Sasson et al., 2017). Thus, selecting animals with less methane production could lead to a decrease in the
289 overall impact of livestock in our environment.

290
291

292 **Conclusions**

293 Illumina sequencing is the most used sequencing technology for shotgun metagenomics as it has a high
294 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](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-](https://doi.org/10.1111/1751-7915.12141)
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

- 332 communication: Signs of host genetic regulation in the microbiome composition in 2 dairy breeds:
333 Holstein and Brown Swiss. *Journal of Dairy Science*, 101(3), 2285–2292.
334 <https://doi.org/10.3168/jds.2017-13179>
- 335 Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation
336 sequencing technologies. *Nature Reviews Genetics*, 17(6), 333–351. <https://doi.org/10.1038/nrg.2016.49>
- 337 Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., ... Chiu, C. Y. (2015). Rapid
338 metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing
339 analysis. *Genome Medicine*. <https://doi.org/10.1186/s13073-015-0220-9>
- 340 Huson, D. H., Albrecht, B., Bağci, C., Bessarab, I., Górska, A., Jolic, D., & Williams, R. B. H. (2018).
341 MEGAN-LR: New algorithms allow accurate binning and easy interactive exploration of metagenomic
342 long reads and contigs. *Biology Direct*. <https://doi.org/10.1186/s13062-018-0208-7>
- 343 Jami, E., White, B. A., & Mizrahi, I. (2014). Potential role of the bovine rumen microbiome in modulating
344 milk composition and feed efficiency. *PLoS ONE*, 9(1). <https://doi.org/10.1371/journal.pone.0085423>
- 345 Jewell, K. A., McCormick, C. A., Odt, C. L., Weimer, P. J., & Suen, G. (2015). Ruminant bacterial community
346 composition in dairy cows is dynamic over the course of two lactations and correlates with feed
347 efficiency. *Applied and Environmental Microbiology*. <https://doi.org/10.1128/AEM.00720-15>
- 348 Kelly, W. J., Pacheco, D. M., Li, D., Attwood, G. T., Altermann, E., & Leahy, S. C. (2016). The complete
349 genome sequence of the rumen methanogen *Methanobrevibacter millerae* SM9. *Standards in Genomic
350 Sciences*. <https://doi.org/10.1186/s40793-016-0171-9>
- 351 Knapp, J. R., Laur, G. L., Vadas, P. A., Weiss, W. P., & Tricarico, J. M. (2014). Invited review: Enteric
352 methane in dairy cattle production: Quantifying the opportunities and impact of reducing emissions.
353 *Journal of Dairy Science*. <https://doi.org/10.3168/jds.2013-7234>
- 354 Koliada, A., Syzenko, G., Moseiko, V., Budovska, L., Puchkov, K., Perederiy, V., ... Vaiserman, A. (2017).
355 Association between body mass index and Firmicutes/Bacteroidetes ratio in an adult Ukrainian
356 population. *BMC Microbiology*. <https://doi.org/10.1186/s12866-017-1027-1>
- 357 Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: scalable
358 and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*,
359 27(5), 722–736. <https://doi.org/10.1101/gr.215087.116>
- 360 Leggett, R. M., Alcon-Giner, C., Heavens, D., Caim, S., Brook, T. C., Kujawska, M., ... Clark, M. D. (2017).
361 Rapid MinION metagenomic profiling of the preterm infant gut microbiota to aid in pathogen
362 diagnostics. *BioRxiv*, 180406. <https://doi.org/10.1101/180406>
- 363 Leo Lahti and Sudarshan Shetty. (n.d.). *microbiome R package*. Retrieved from <http://microbiome.github.io>
- 364 Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: An ultra-fast single-node
365 solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*,
366 31(10), 1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>
- 367 Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly.
368 *Genomics, Proteomics and Bioinformatics*. <https://doi.org/10.1016/j.gpb.2016.05.004>

- 369 Ma, X., Stachler, E., & Bibby, K. (2017). Evaluation of Oxford Nanopore MinION Sequencing for 16S rRNA
370 Microbiome Characterization. *BioRxiv*, 1–19. <https://doi.org/10.1101/099960>
- 371 Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome
372 assemblies. *Bioinformatics*, 27(21), 2957–2963. Retrieved from
373 <http://dx.doi.org/10.1093/bioinformatics/btr507>
- 374 McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R Package for Reproducible Interactive Analysis and
375 Graphics of Microbiome Census Data. *PLoS ONE*, 8(4). <https://doi.org/10.1371/journal.pone.0061217>
- 376 Mikheenko, A., Valin, G., Prjibelski, A., Saveliev, V., & Gurevich, A. (2016). Icarus: Visualizer for de novo
377 assembly evaluation. *Bioinformatics*, 32(21), 3321–3323. <https://doi.org/10.1093/bioinformatics/btw379>
- 378 Newbold, C. J., De la Fuente, G., Belanche, A., Ramos-Morales, E., & McEwan, N. R. (2015). The role of
379 ciliate protozoa in the rumen. *Frontiers in Microbiology*. <https://doi.org/10.3389/fmicb.2015.01313>
- 380 Oksanen, A. J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., ... Szoecs, E. (2019).
381 *Package 'vegan.'*
- 382 Poulsen, M., Schwab, C., Borg Jensen, B., Engberg, R. M., Spang, A., Canibe, N., ... Urich, T. (2013).
383 Methylotrophic methanogenic Thermoplasmata implicated in reduced methane emissions from bovine
384 rumen. *Nature Communications*. <https://doi.org/10.1038/ncomms2432>
- 385 Sasson, G., Kruger Ben-Shabat, S., Seroussi, E., Doron-Faigenboim, A., Shterzer, N., Yaacoby, S., ...
386 Mizrahi, I. (2017). Heritable Bovine Rumen Bacteria Are Phylogenetically Related and Correlated with
387 the Cow's Capacity To Harvest Energy from Its Feed. *MBio*, 8(4). Retrieved from
388 <http://mbio.asm.org/content/8/4/e00703-17.abstract>
- 389 Schmieder, R., & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets.
390 *Bioinformatics*, 27(6), 863–864. Retrieved from <http://dx.doi.org/10.1093/bioinformatics/btr026>
- 391 Seshadri, R., Leahy, S. C., Attwood, G. T., Teh, K. H., Lambie, S. C., Cookson, A. L., ... Cerón Cucchi, M.
392 (2018). Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection.
393 *Nature Biotechnology*. <https://doi.org/10.1038/nbt.4110>
- 394 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
395 microbiome using full-length 16S rRNA amplicon sequencing. *Scientific Reports*.
396 <https://doi.org/10.1038/srep29681>
- 397 Sims, D., Sudbery, I., Ilott, N. E., Heger, A., & Ponting, C. P. (2014). Sequencing depth and coverage: Key
398 considerations in genomic analyses. *Nature Reviews Genetics*, 15(2), 121–132.
399 <https://doi.org/10.1038/nrg3642>
- 400 Tan, M. H., Austin, C. M., Hammer, M. P., Lee, Y. P., Croft, L. J., & Gan, H. M. (2018). Finding Nemo:
401 Hybrid assembly with Oxford Nanopore and Illumina reads greatly improves the clownfish (*Amphiprion*
402 *ocellaris*) genome assembly. *GigaScience*, 7(3), 1–6. <https://doi.org/10.1093/gigascience/gix137>
- 403 Tapio, I., Snelling, T. J., Strozzi, F., & Wallace, R. J. (2017). The ruminal microbiome associated with
404 methane emissions from ruminant livestock. *Journal of Animal Science and Biotechnology*, 8(1), 1–11.
405 <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
<i>Species</i>	125	146
<i>Genera</i>	78	86
<i>Phyla</i>	17	21
<i>Read length (pb)</i>	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

419

420

421 **Table 3. Assembly statistics**

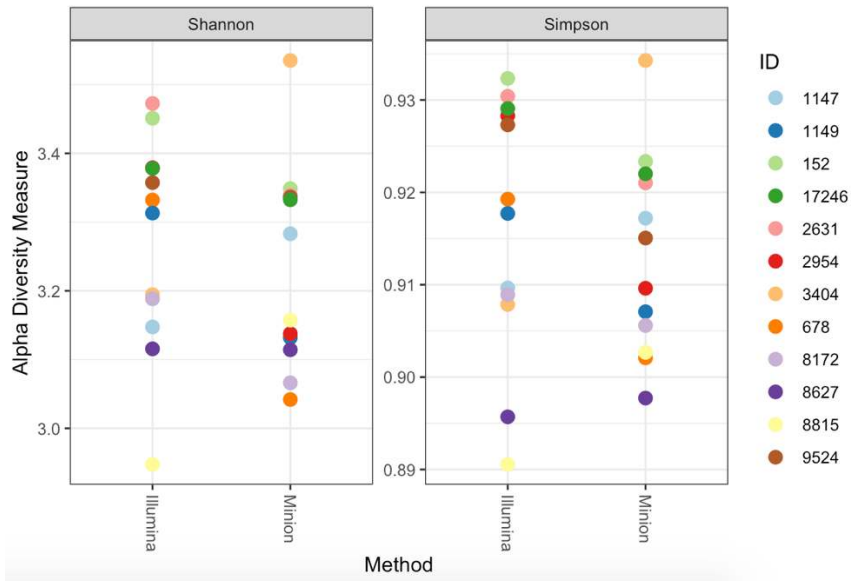
	Illumina assembly	MinION assembly
Length (bp)	5,296,436	2,550,082
GC (%)	47.7	46.89
N's	0	0
N50	641	11,549
Largest Contig (bp)	11,173	49,178
Contigs	7,856	281

422

423

424

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

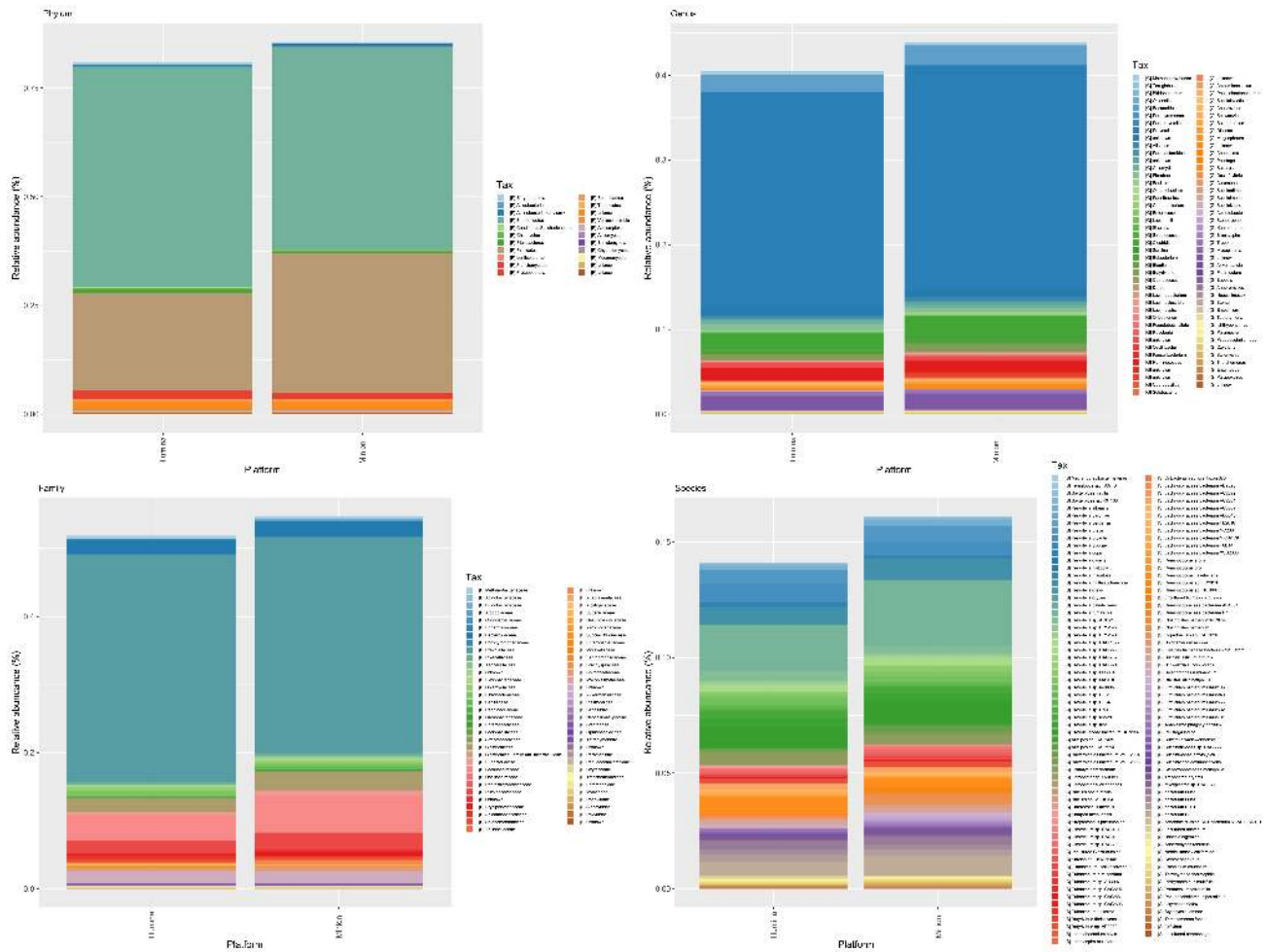
428

429

430

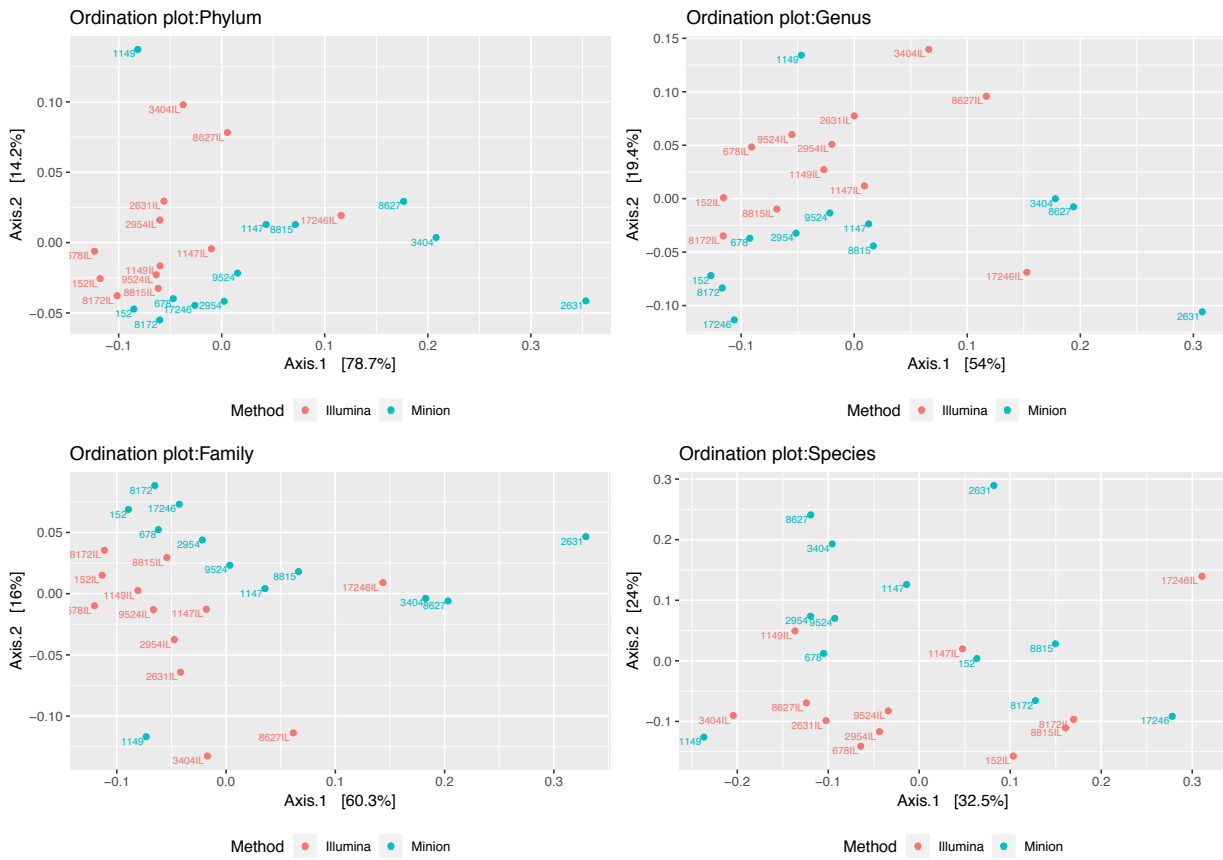
431

432 **Figure 1. Comparison of relative abundances of common taxa at four different taxonomic levels using Illumina and**
433 **MinION sequencers. Relative abundances are similar for common taxa at all levels, although the phylum Firmicutes has**
434 **a higher relative abundance in the MinION dataset which counts for most of the observed differences in the four**
435 **graphics.**
436



437
438

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



444

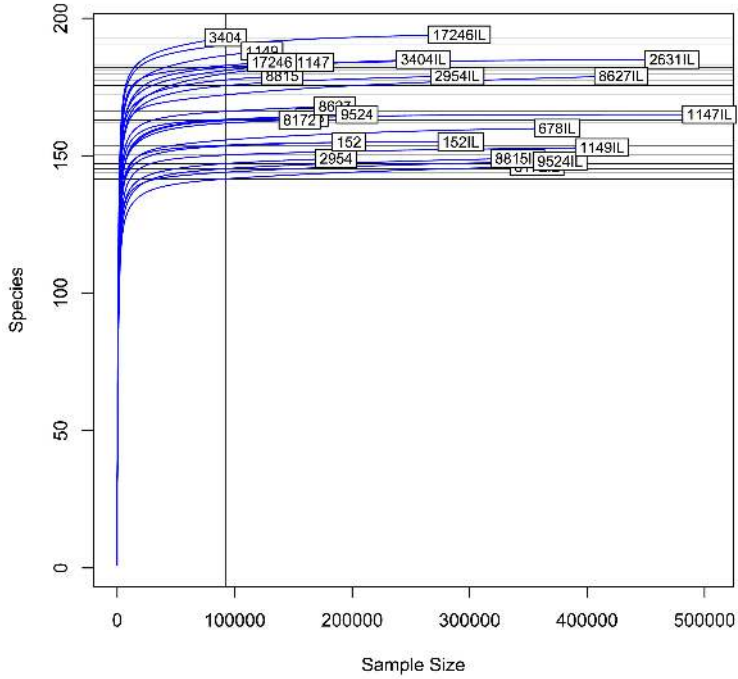
445

446

447 **Supplementary Figures**

448

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

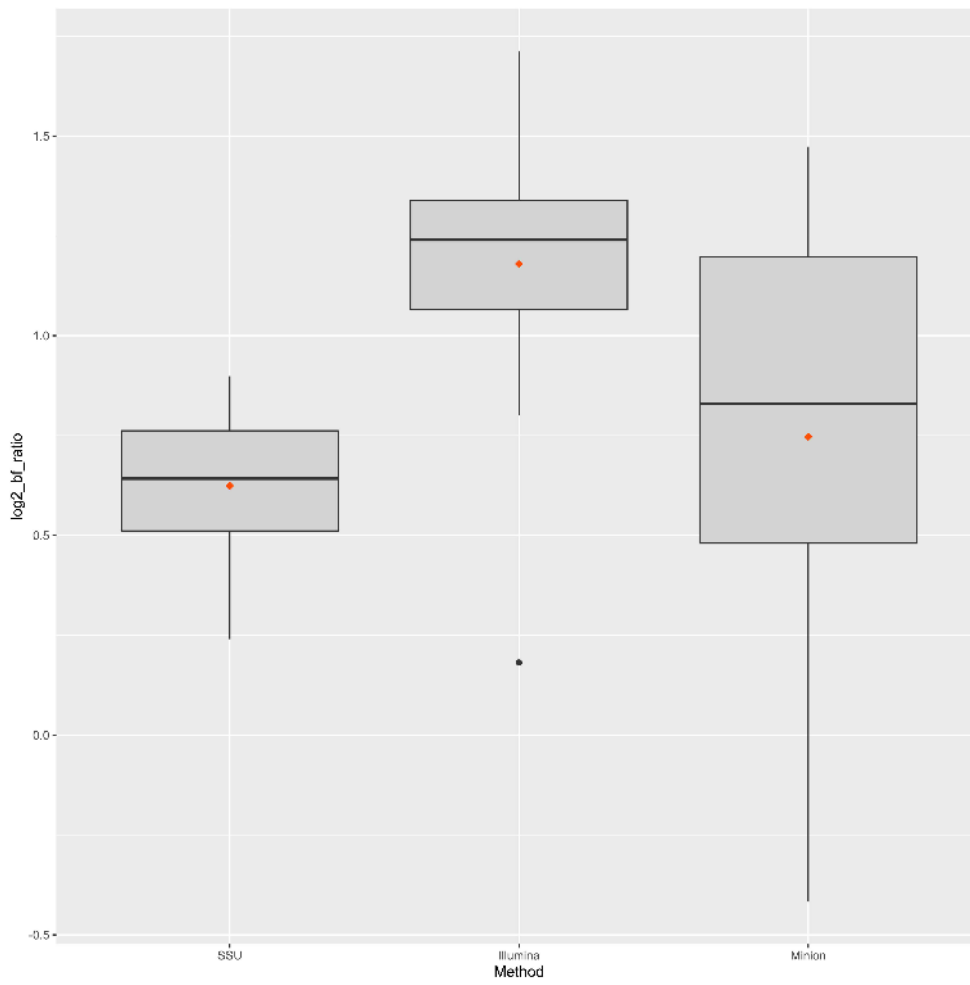


450

451

452

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



454