Phylogenetic diversity and dietary association of rumen *Treponema* revealed using group-specific 16S rRNA gene -based analysis

Aschalew Z. Bekele, Satoshi Koike* and Yasuo Kobayashi

Graduate School of Agriculture, Hokkaido University, Sapporo, Japan

***Correspondence:** Satoshi Koike, Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan.

Tel: +81 (11) 706-2812, Fax: +81 (11) 706-2814.

E-mail: skoike7@anim.agr.hokudai.ac.jp

Key words: rumen, *Treponema*, diversity, diet association, 16S rRNA gene

Running title: Phylogenetic diversity and diet association of rumen *Treponema*

Abstract

Treponema spp. are a commonly detected bacterial group in the rumen that are involved in the degradation of soluble fibers. In this study, a ruminal *Treponema* group-specific PCR primer targeting the 16S rRNA gene was designed and used to assess the phylogenetic diversity and diet association of this group in sheep rumen. Total DNA was extracted from rumen digesta of three sheep fed a diet based on alfalfa/orchardgrass hay or concentrate. The real-time PCR quantification indicated that the relative abundance of the *Treponema* group in the total rumen bacteria was as high as 1.05%, while the known species *Treponema bryantii* accounted for only 0.02%. Fingerprints of the *Treponema* community determined by 16S rDNA targeted denaturing gradient gel electrophoresis (DGGE) analysis tended to differ among the diets. Principal component analysis of the DGGE profiles distinguished those *Treponema* associated with either the hay or the concentrate diets. Analysis of a *Treponema* 16S rRNA gene clone library showed phylogenetically distinct operational taxonomic units for a specific dietary condition, and significant ($P = 0.001$) differences in community composition were observed among clone libraries constructed from each dietary regimen. The majority of clones (75.4%) had less than 97% sequence similarity with known *Treponema*. These results suggest the predominance of uncultured *Treponema* that appear to have distinct members related to the digestion of either hay or concentrate diet.

Introduction

The distribution of spiral-shaped bacteria (spirochetes) and their role in the degradation of plant material in the rumen have been reported by different workers (Bryant, 1952; Stanton & Canale-Parola, 1979; Ziolecki & Wojciechowicz, 1980). Direct microscopic enumeration of spirochetes showed up to 2.0×10^8 cells/ml of bovine rumen fluid (Stanton & Canale-Parola, 1979), which is comparable to the population density of common rumen bacterial species (Bryant & Burkey, 1953). All strains of spirochetes isolated from the rumen have been classified in the genus *Treponema*, comprised of three described species: *Treponema bryantii* (Stanton & Canale-Parola, 1980), *Treponema saccharophilum* (Paster & Canale-Parola, 1985) and *Treponema zioleckii* (Piknova *et al*., 2008). Rumen *Treponema* strains are able to degrade plant polysaccharides (Ziolecki, 1979), and *in vitro* studies have shown a beneficial interaction of *T. bryantii* with the cellulolytic bacterium *Fibrobacter succinogenes* (Stanton & Canale-Parola, 1980).

Recent application of molecular techniques in the study of microbial ecology demonstrated the existence of a considerable proportion of diverse uncultivated spirochetes involved in chronic disease in the human oral cavity and in degradation of lignocellulose materials in the termite gut (Paster *et al*., 1996 & 2001; Dewhirst *et al*., 2000). For example, 16S rRNA gene-based clone library analysis of samples from the oral cavity of a human subject and from the hindgut of a single termite species respectively suggested some 20 and 23 new species of spirochetes (Choi *et al*., 1994; Lilburn *et al*., 1999). Considering the individuality of human microbiota and the existence of ~280 termite genera, these observations suggest the presence of a great diversity of spirochetes, particularly uncultured members. In contrast to the above digestive tract environments, our knowledge of the uncultured *Treponema* community in the rumen is very limited. Current understanding of the rumen *Treponema* diversity is mainly based on earlier cultivation-based studies that showed

morphological and physiological variation in rumen spirochetes (Paster *et al*., 1991, Piknova *et al.*, 2008). A comprehensive analysis of 16S rRNA gene sequences derived from the rumen showed that rumen *Treponema* were not frequently detected (Edwards *et al*., 2004; Yang *et al.,* 2010).

However, we had previously retrieved a number of *Treponema* clones related to both cultured and uncultured members from a fiber-associated community (Koike *et al*., 2003; Shinkai *et al*., 2010). Based on these data, we speculated that rumen *Treponema* diversity has been underestimated and members of this group may play a metabolic role in fiber degradation. In this study, a ruminal *Treponema* group-specific PCR primer was designed and used to determine the population size, phylogenetic diversity and distribution of the *Treponema* community in the rumen. By comparing 16S rRNA gene sequences from sheep fed different diets, we tested the hypothesis that distinct members of *Treponema* may relate to the digestion of either hay or concentrate diet.

Materials and methods

Animals and sampling

All procedures with live animals were approved by the Animal Care and Welfare Committee of Hokkaido University, Japan (Protocol number 09-0046). Three rumen fistulated sheep (average body weight, 90.7 ± 6.9 kg) were used in three consecutive periods corresponding to three dietary regimens. In the first period each animal was given an alfalfa hay diet (1.2 kg/day), and in the second period an orchardgrass hay diet (1.2 kg/day). The orchardgrass hay diet was supplemented with soybean meal. In the third period, each animal was fed a concentrate-diet containing 1.0 kg of a commercial formula feed (Ram 76ME, Mercian, Tokyo, Japan) and 0.5 kg of the orchardgrass hay. The three diets were formulated to be isonitrogenous (18.2% crude protein). Each diet was fed for 3 weeks and rumen contents were sampled from individual animals prior to feeding on the last day of the experimental

period. The samples were stored at -30°C until DNA was extracted. Throughout the experimental period, animals were kept in the individual pens and fed once daily at 09:00 h. Water and a mineral block was available *ad libitum*.

DNA extraction

Total DNA was extracted from 0.25 g wet rumen content samples following the RBB+C method according to Yu and Morrison (2004). Briefly, cells were lysed by repeated beating with glass beads (Mini Bead Beater, BioSpec Products, Bartlesville, OK, USA) in the presence of 4% (w/v) sodium dodecyl sulfate, 500 mM NaCl, 50 mM Tris-HCl (pH 8.0) and 50 mM EDTA. Two different sized (0.1 mm and 0.5 mm) glass beads were used for disrupting the cells. After incubation of the lysate at 70°C for 15 min, nucleic acids were recovered by isopropanol precipitation. DNA was treated with DNase-free RNase and proteinase K, and purified by a QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany). Purified DNA was quantified by a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA) and the final concentration of DNA extracts was adjusted to 10 ng/µl for use in all downstream applications.

Design of PCR primer

In order to design a PCR primer targeting rumen *Treponema*, 63 currently available 16S rRNA gene sequences of rumen *Treponema* were obtained from the GenBank database as well as from our clone library sequence collections. Sequences for the three known *Treponema* species were also included in the analysis. In addition, 10 mammalian and 14 termite *Treponema* sequences were included in the *in silico* analysis. The sequences were aligned with Clustal X v.1.81 multiple sequence alignment software (Thompson *et al*., 1997). The *Treponema* group-specific forward primer was designed based on a region conserved among all rumen *Treponema,* while the universal primer 926R (Watanabe *et al*., 2001) was chosen as a reverse primer. The nucleotide positions of the target site for the forward primer

126 on *T. bryantii* 16S rRNA gene sequences were 380 to 400 while those of the reverse primer
127 were 934 to 953, yielding a 575 bp PCR product. The primer set was designed to cover all
128 rumen *Treponema* and named g-TrepoF. The online BLAST program
129 (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to determine the specificity of the forward
130 primer. The specificity of the primers was further tested by PCR amplification using genomic
131 DNA from pure cultures of 16 representative rumen bacterial strains including *Treponema*
132 *bryantii* ATCC33254, *Fibrobacter succinogenes* ATCC19169, *Ruminococcus albus* 8,
133 *Ruminococcus flavefaciens* C94, *Prevotella ruminicola* 23, *Prevotella bryantii* $B_1$4,
134 *Prevotella brevis* GA33, *Butyrivibrio fibrisolvens* H17c, *Butyrivibrio fibrisolvens* D1,
135 *Eubacterium ruminantium* GA195, *Selenomonas ruminantium* GA192, *Succinivibrio*
136 *dextrinosolvens* ATCC19716, *Succinimonas amylolytica* ATCC19206, *Streptococcus bovis*
137 ATCC33317, *Megasphaera elsdenii* ATCC25940, and *Anaerovibrio lipolytica* ATCC33276.
138 Rumen *Treponema* group-specific clone libraries constructed using the primers also served to
139 confirm primer specificity. The sequences of all primers used in this study are shown in Table
140 1.

141 **Real-time PCR quantification of the 16S rRNA gene**

142 Plasmid DNA to be used as the standard in real-time PCR was obtained by cloning of 16S
143 rRNA gene PCR products into *Escherichia coli* JM109 cells, as previously described (Koike
144 *et al.*, 2007). For *Treponema* group-specific PCR as well as *T. bryantii*-specific PCR, a 16S
145 rRNA gene fragment of *T. bryantii* ATCC33254 was used to prepare a plasmid DNA
146 standard as previously reported (Bekele *et al.,* 2010). The PCR primers used are shown in
147 Table 1. PCR amplification for the quantification of target bacterial 16S rRNA gene was
148 performed with a LightCycler 2.0 system (Roche Applied Science, Penzberg, Germany) and
149 FastStart DNA Master SYBR Green I (Roche Applied Science, Penzberg, Germany). The
150 optimal amplification conditions for each primer pair were achieved with 3.5 mM $MgCl_2$.

151 The 20 µl reaction mixture contained 2.5 mM $MgCl_2$, 2 µl 10× Mastermix (containing
152 FastStart *Taq* DNA polymerase, reaction buffer, dNTP mixture, 1 mM $MgCl_2$ and SYBR
153 Green I dye), 0.5 µM of each primer and 10 ng template DNA. The thermal profile consisted
154 of denaturation at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, annealing at the
155 temperature indicated for the primer pair (Table 1) for 5 s and 72°C for an appropriate
156 extension time (Table 1). Dissociation curve analysis was performed to ascertain the
157 specificity of amplicons by slow heating with a 0.1°C/s increment from 70°C to 95°C, with
158 fluorescence collection at 0.1°C intervals. A 10-fold dilution series of the plasmid DNA
159 standard for the respective target bacterial 16S rRNA gene was run along with the samples.
160 The respective genes were quantified using standard curves obtained from the amplification
161 profile of known concentrations of the plasmid DNA standard. To obtain the relative
162 abundance of *Treponema* in the rumen, the assay values for 16S rRNA genes of the target
163 group or species were normalized to the total number of copies of rumen bacterial 16S rRNA
164 genes. The normalized assay values were analyzed statistically by single factor ANOVA at a
165 level of significance of 0.05.

166 ***Treponema* specific 16S rDNA denaturing gradient gel electrophoresis**

167 Denaturing gradient gel electrophoresis (DGGE) was used to examine the relationship
168 between diet and the rumen *Treponema* community. The analysis was carried out in a Bio-
169 Rad DCode Universal Mutation Detection System (Bio-Rad Laboratories, Hercules, CA,
170 USA). The g-TrepoF and BAC926R primers employed for real-time PCR were used to
171 amplify the V3-V5 regions of the 16S rRNA gene of *Treponema* in the sheep rumen samples.
172 Genomic DNA from *T. bryantii* ATCC 33254 was also included in the analysis. An amplicon
173 of ca. 575 bp for DGGE analysis was obtained by modifying the reverse primer by addition
174 of a 40 bp GC clamp
175 (5'CGCCCGCCGCGCGCGGCGGGCGGGGCGGGGGCACGGGGGG-3'). PCR was

conducted with a Veriti 96-well thermal cycler (Applied Biosystems, Singapore). A reaction mixture containing 0.4 µM of each primer, 5 µl of 10× ExTaq buffer, 0.2 µM of each dNTP, 1.25 U ExTaq polymerase (Takara, Otsu, Japan), and 10 ng of template DNA in a total volume of 50 µl was prepared. The temperature program for cycling consisted of an initial denaturation at 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, annealing at 64 °C for 30 s, and extension at 72°C for 30 s with a final extension at 72°C for 5 min. PCR-amplified 16S rRNA gene fragments were separated using an 8% polyacrylamide gel with 0.5× TAE buffer (20 mM Tris-acetate, 10 mM sodium acetate, 0.5 mM EDTA, pH 8.0) and a 35% to 60% linear gradient of denaturant (100% denaturant corresponded to 40% (v/v) deionized formamide and 7 M urea). Each gel was run at 60°C, 80 V for 16 h, and then placed in fixing solution (10% ethanol and 0.5% acetic acid) for 2 h, stained in 0.1% (w/v) silver nitrate solution for 20 min and developed in 1.5% sodium hydroxide (w/v), 0.1% sodium borohydride (w/v) and 0.4% formaldehyde (v/v) for 8 min. Thereafter, the gel was rinsed and kept in distilled water until the image was scanned. Gel images were analyzed by BioNumerics software version 4.5 (Applied Maths, Kortrijk, Belgium). Normalized banding patterns were used to generate dendrograms by calculating Dice similarity coefficients and by an unweighted pair group method with an arithmetic averages clustering algorithm. For statistical analysis, the DGGE banding patterns were converted into binary data as presence or absence of bands using Bionumerics software and principal component analysis (PCA) was conducted using the PRIMER 5 data analysis software system (PRIMER-E Ltd, Plymouth, UK).

16S rRNA gene clone library construction and sequencing

Three clone libraries were constructed for the respective feeding conditions. Mixed DNA samples obtained from the rumen content DNA from three animals under the same dietary conditions were used for library construction. PCR products were generated by the primers g-

TrepoF and BAC926R with the same reaction and amplification conditions described for DGGE with the exception of the reverse primer without GC clamp. PCR products were cloned with a pGEM-T Easy Vector System (Promega, San Luis Obispo, CA, USA) according to the manufacturer's instructions. Clones containing the correct insert were sequenced at Takara Bio (Yokkaichi, Japan). Clone nomenclature was as follows: for the alfalfa and orchardgrass hay-associated *Treponema* libraries, clone names began with ALTC and OGTC, respectively, followed by the clone number. Clone names in the concentrate-associated *Treponema* library began with CTC followed by the clone number. All the sequences were deposited into the GenBank database with the accession numbers AB537568 through AB537880.

Sequence analysis

A total of 313 16S rRNA gene sequences, obtained from the three clone libraries and representative rumen *Treponema* sequences from the NCBI database, were included in the analysis. The sequences were automatically aligned using ClustalX ver.1.81 multiple sequence alignment software (Thompson *et al.*, 1997). A neighbor-joining tree (Saito & Nei, 1987) with a Kimura-2 correction was constructed in MEGA v.3.1 software. (Tamura *et al.,* 2007). In order to statistically evaluate the branching of the tree, bootstrap analysis (Felsenstein, 1985) was carried out with 1000 resamplings of the data. Sequences from the three rumen *Treponema* clone libraries were compared with 16S rRNA gene sequences in the GenBank database using the Basic Local Alignment Search Tool (BLAST program, Altschul *et al.*, 1990) to obtain similarity values.

Operational taxonomic units (OTUs) were defined based on a 97% sequence identity criterion (Stackebrandt & Goebel, 1994). Analysis of the diversity for the individual and combined libraries was carried out using the nonparametric estimator Chao1 (Chao, 1984) and the Shannon index (Shannon & Weaver, 1949) using FastGroupII software.

(http://biome.sdsu.edu/fastgroup/fg_tools.htm). The percentage of coverage of the clone libraries was calculated by Good's method with the formula [1 - (n/N)] × 100, where n is the number of singletons and N is the total number of sequences (Good, 1953). The statistical differences among the 16S rRNA gene clone libraries from the respective feeding conditions were compared using the web-based Library Shuffling (web-LIBSHUFF) program version 0.96 (http/libshuff.mib.uga.edu) (Henriksen, 2004) to determine whether a given pair of the libraries was drawn from the same population. The significant difference level for comparison of the three libraries was defined as $P = 0.0085$. The sequences were initially aligned by ClustalX and genetic distances were generated in the DNADIST program of the PHYLIP package (v.3.67) (Felsenstein, 2007) using the Jukes–Cantor model before submitting to web-LIBSHUFF.

RESULTS

Primer specificity and validation

The newly designed g-TrepoF primer showed a 100% sequence match with all ruminal (n=63) and other mammalian (n=10) *Treponema* sequences tested, while 50% of the tested termite (n=14) *Treponema* sequences had a single nucleotide mismatch in the middle of the primer site (data not shown). Therefore, it was assumed that the g-TrepoF primer covers all rumen *Treponema* and also has a broad coverage of non-ruminal *Treponema*. The specificity of the primer (g-TrepoF) for rumen *Treponema* was also validated by using an online BLAST similarity search and by PCR amplification of 16 representative rumen bacteria. The BLAST similarity search of the primer sequences showed similarity with 16S rRNA gene sequences of spirochetes. The primer set g-TrepoF and BAC926R did not cross-react with any of the non-target rumen bacteria tested at the specified PCR conditions, while PCR products of the expected size were obtained from *T. bryantii* genomic DNA (data not shown). The *Treponema* clone libraries constructed from DNA extracts of rumen digesta of sheep also

confirmed the specificity of the primers for rumen *Treponema*. No other bacterial 16S rRNA gene sequences other than *Treponema* were detected in the libraries.

Although primer sets that yield short amplicons are ideal for real time PCR amplification, it was difficult to design primers that are specific for *Treponema* and yield a smaller PCR product. The g-TrepoF and the BAC926R primer set yield a relatively large (575bp) PCR product. However, the standard curve for the assay was comparable with those of the total bacterial and *T. bryantii* species-specific primers producing PCR efficiencies higher than 1.9 (Table 1).The dissociation curve obtained for the samples had a similar melting point with the standard plasmid DNA indicating that there were no non-specific amplifications. The g-TrepoF and BAC926 primers produced a single dissociation curve peak at 90 °C when tested against DNA from *T. bryantii* and when using total rumen microbial DNA.

Relative abundance of *Treponema* in the rumen

The relative proportions of the 16S rRNA gene copies for the *Treponema* group and *T. bryantii* are shown in Table 2. The mean relative population size of the *Treponema* group in the total rumen bacteria of sheep fed alfalfa diet was as high as 1.05%, while that of *T. bryantii* was only 0.02%. Although the highest population size of *Treponema* was found in the alfalfa-fed sheep, diet did not significantly affect the *Treponema* group ($P = 0.648$) or the *T. bryantii* ($P = 0.977$) population.

DGGE banding patterns of rumen *Treponema*

The DNA fingerprints of *T. bryantii* showed a single band, while a number of bands were observed for the other *Treponema* in the rumen content DNA samples from sheep fed different diets. The DGGE profiles of the *Treponema* community associated with the hay (alfalfa and orchardgrass) and concentrate diets showed different banding patterns. The DGGE profiles across diet showed consistently fewer bands (except animal 3) in samples

from concentrate-fed animals (Fig. 1). The PCA of the binary data of DGGE profiles distinguished *Treponema* population that associated with either the hay or the concentrate diets resulting in two clusters (Fig. 2), although one exception was observed.

Sequence diversity and community composition

Based on a 97% sequence similarity criterion, the 313 clone sequences from the combined libraries were grouped into 67 OTUs (Table 3). A similar number of OTUs (30 to 32) was identified for each diet. Good's coverage of the combined library was 91.1%, while the coverage for the alfalfa, orchardgrass and concentrate libraries was 83.8%, 88.1% and 85.2%, respectively (Table 3). Although the Chao1 estimation was lower for the orchardgrass, the predicted OTUs and the overall level of diversity estimation by the Shannon index were higher for the alfalfa and orchardgrass hay libraries (Table 3), which correlated with the DGGE observation (Fig. 1).

Among the 77 (24.6%, 2 OTUs) clone sequences that showed 97% or more sequence similarity with cultured *Treponema,* 76 were related to *T. bryantii.* Only a single sequence related to *T. zioleckii* and no sequences having 97% or more similarity with *T. saccharophilum* were found. The majority of clones (236 clones, 75.4%) were related to uncultured *Treponema*, irrespective of diet (Table 3). Among the uncultured *Treponema*, 70 clones had 97% or more similarity with sequences of uncultured *Treponema* clones, while 166 clones showed 86 to 96% similarity (Table 3) with any sequence in the NCBI database. Pairwise comparison of each 16S rRNA gene library using web-LIBSHUFF confirmed that the libraries were significantly ($P = 0.001$) different from one another (data not shown).

The results of a phylogenetic analysis of the 67 OTUs identified among the combined 16S rRNA gene sequences from the three libraries are shown in Fig. 3. The phylogenetic tree (Fig. 3) was divided into two major clades (Clade I and II). Additionally, clade II was further categorized in to sub-clades (a-e), although this was not supported by higher bootstrap values.

301 The distribution of clones in the different clades was shown by pie charts with the size of the
302 pie charts corresponding to the size of the clones in each clade. In clade I, 59 clones (58.4%)
303 were from the concentrate library, while in clade II 185 clones (87.3%) were from the hay
304 libraries.

305 ## DISCUSSION

306 16S rRNA gene-based clone libraries constructed using universal PCR primers have been
307 used to monitor the entire rumen bacterial community (Whitford *et al.*, 1998; Tajima *et al.*,
308 1999; Koike *et al.*, 2003; Sundset *et al*., 2007). However, such universal libraries do not
309 sufficiently represent the diversity of specific groups of bacteria in a complex gut
310 environment (Li *et al*., 2008). Our recent analysis of the rumen *Prevotella* community based
311 on group-specific clone libraries showed the abundance of novel rumen *Prevotella* previously
312 undetected (Bekele *et al*., 2010), indicating the advantage of this approach. In the present
313 study, we focused on *Treponema*, a frequently detected rumen bacterial group that has been
314 implicated in the degradation of fiber (Koike et al., 2003; Shinkai et *al*., 2010). A *Treponema*
315 group-specific primer was successfully developed and used to illustrate the diversity and
316 molecular ecology of rumen *Treponema*.

317 Real-time PCR quantification revealed that the relative abundance of *Treponema* was
318 comparable to or higher than that of the other representative rumen bacteria (Stevenson and
319 Weimer, 2007; Bekele *et al*., 2010). Therefore, the *Treponema* group may be one of the core
320 members of the rumen bacterial community. The proportion of *T. bryantii* was about 2% in
321 the *Treponema* group (0.02% vs. 1.05%), indicating that the uncultured *Treponema* were
322 more abundant than cultured representatives. Analysis of the *Treponema* 16S rRNA gene
323 libraries supported this finding (see below). Although a single sequence was identified as *T.*
324 *zioleckii* in the present study, no 16S rRNA gene sequence having 97% or more similarity
325 with *T. sacchrophilum* and *T. zioleckii* was reported in previous studies (Whitford *et al*.,

326 1998; Tajima *et al*., 1999; Koike *et al*., 2003). Therefore, *T. sacchrophilum* and *T. zioleckii*
327 appear to be minor bacterial species in the rumen.

328 Sequence analysis of 16S rRNA gene clone libraries constructed in this study for
329 rumen *Treponema* revealed the presence of phylogenetically diverse and previously
330 undetected OTUs of the rumen *Treponema* community. The DGGE data further showed
331 diverse bands in the animals fed alfalfa and orchardgrass hay. This finding corresponded with
332 diversity analysis of the libraries, which showed higher Shannon index diversity values for
333 the hay diets. A plausible explanation for this finding would be that more diverse members of
334 *Treponema* are involved in the degradation of hay diets.

335 Considering the higher percentage (91.1%) of Good's coverage for the combined
336 library, our library was comprehensive and likely represented the majority of *Treponema* in
337 the sheep rumen. It has been suggested that a group-specific clone library approach could
338 identify more diverse members in the target group than a universal library analysis (Hayashi
339 *et al*., 2006). In human gut studies, attempts to recover diverse members of *Bacteroides* spp.
340 by increasing the size of libraries constructed by universal primers did not result in higher
341 diversity of *Bacteroides* (Li et al., 2008). Preferential PCR amplification of certain groups of
342 rumen microbes has been suggested as a possible reason for the difficulty in detecting a
343 particular group with universal primers (Tajima *et al*., 2001), and this may explain the low
344 level detection of *Treponema* sequences in previous studies (Whitford *et al.*, 1998; Tajima *et*
345 *al.*, 1999; Ozutsumi *et al*., 2005). Therefore, the group-specific clone library approach that
346 we followed in this study proved useful to obtain a comprehensive description of the diversity
347 of *Treponema* in the rumen.

348 Phylogenetic analysis of the *Treponema* 16S rRNA gene sequences showed a closer
349 phylogeny of clones retrieved from a particular diet. In the phylogenetic tree, clade I was
350 mainly comprised of clones (58.4% of the overall concentrate clones) associated with

concentrate feeding; while clade II predominantly consisted of *Treponema* clones (87.3% of the overall hay clones) associated with hay feeding. These findings suggest that closely related phylotypes of rumen *Treponema* associate with a given diet, and may play a role in the degradation of that particular diet in the rumen. This interpretation of the phylogenetic analysis was supported by results of the principal component analysis of the DGGE fingerprints of the *Treponema* community that showed separate clusters for *Treponema* associated with either the hay or the concentrate diets.

Pairwise comparison of each 16S rRNA gene library indicated that the composition of *Treponema* associated with the concentrate diet differed from those associated with the hay diets. Similarly, the *Treponema* community associated with each hay diet differed significantly ($P = 0.001$). Therefore, differences observed among the libraries were attributed to the presence of phylotypes specifically associated with a given diet. Several studies have shown that some ruminal bacterial species are indeed very specialized; while specialist's vs. generalist's others have a broad range of substrate specificity (Krause and Russell, 1996). Diet-dependent shifts in the entire bacterial community have also been interpreted as changes caused by the specialized niches and substrate requirements of different rumen bacteria (Tajima *et al*., 2001, Welkie *et al*., 2010). Recently, we reported molecular evidence for the existence of diet-specific subpopulations of *Prevotella* that might be involved in the degradation of either hay or concentrate diets (Bekele *et al*., 2010). Collectively, these findings support the concept of functional specialization among rumen bacterial groups and even within a bacterial group such as *Treponema.*

Two OTUs (25 and 67) had a phylogenetic position closer to cultured species of *T. bryantii* and *T. saccharophilum,* respectively. These OTUs may have similar functions to that of the cultured close relatives. Cultured rumen *Treponema* strains do not break down cellulose, but are capable of catabolizing other structural polysaccharides such as pectin,

xylan, and fructan (Ziolecki, 1979; Wojciechowicz & Ziolecki, 1979; Ziolecki & Wojciechowicz, 1980; Piknova *et al*., 2008), and also of utilizing hydrolysis products of plant polymers such as cellobiose, xylose, arabinose and galacturonic acid (Paster & Canale-Parola, 1985). Interestingly, the majority of clones belonging to OTUs 25 and 67 were obtained from the animals fed a hay diet. Therefore, these clones may be involved in rumen fiber degradation.

In conclusion, this study revealed the phylogenetic diversity of rumen *Treponema* in sheep rumen. The population size of ruminal *Treponema* was comparable to that of other representative ruminal species; however, the majority of the members of this group remain uncultured. The diet association of *Treponema* clones suggests the specialized metabolic niches of rumen treponemes related to the digestion of either a hay or concentrate diet. We demonstrated that a group-specific clone library approach can be a useful method to comprehensively profile the composition of specific target bacteria in the rumen, particularly of uncultured members. This study highlights the need for detailed profiling of the huge uncultured component of the rumen bacterial community in order to understand their role in the degradation of feed in the rumen.

Acknowledgements

The authors acknowledge Prof. R.I. Mackie for his helpful suggestions and proof reading of the manuscript.

References

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**:403-410.

Bekele ZA, Koike S & Kobayashi Y (2010) Genetic diversity and diet specificity of ruminal *Prevotella* revealed by 16S rRNA gene-based analysis. *FEMS Microbiol Lett* **305:** 49-57.

400 Bryant MP (1952) The isolation and characteristics of a spirochete from the bovine rumen. *J*
401 *Bacteriol* **64**:325-335
402 Bryant MP & Burkey LA (1953) Numbers and some predominant groups of bacteria in the
403 rumen of cows fed different rations. *J Dairy Sci* **36**:218-224.
404 Chao A (1984) Non-parametric estimation of the number of classes in a population. *Scand J*
405 *Stat* **11:**265-270.
406 Chao A (1987) Estimating the population size for capture-recapture data with unequal
407 catchability. *Biometrics* **43**:783-91.
408 Choi BK, Paster BJ, Dewhirst FE & Gobel UB **(**1994) Diversity of cultivable and
409 uncultivable oral spirochetes from a patient with severe destructive periodontitis.
410 *Infect Immun* **62:**1889–1895.
411 Dewhirst FE, Tamer MA, Ericson RE, Lau CN, Levanos VA, Boches SK, Galvin JL & Paster
412 BJ (2000) The diversity of periodontal spirochetes by 16S rRNA analysis. *Oral*
413 *Microbiol. Immunol* **15:**196–202.
414 Edwards JE, McEwan NR, Anthony JT & Wallace RJ (2004) 16S rDNA library-based
415 analysis of ruminal bacterial diversity. *Antonie van Leeuwenhoek* **86**:263-281.
416 Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap.
417 *Evolution* **39**:783-791.
418 Felsenstein J (2007) PHYLIP-Phylogeny Inference Package (version 3.67), Distributed by the
419 author, Department of Genome Sciences, University of Washington, Seattle.
420 http://evolution.genetics.washington.edu/phylip.html.
421 Good IJ (1953) The population frequencies of species and the estimation of population
422 parameters. *Biometrika* **40**: 237–264.
423 Hayashi H, Sakamoto M, Kitahara M & Benno Y (2006) Diversity of the *Clostridium*
424 *coccoides* group in human fecal microbiota as determined by 16S rRNA gene library.
425 *FEMS Microbiol Lett* **257:**202-207.

426 Henriksen JR 2004. webLIBSHUFF (http://libshuff.mib.uga.edu).

427 Koike S, Yoshitani S, Kobayashi Y & Tanaka K (2003) Phylogenetic analysis of fiber-
428 associated rumen bacterial community and PCR detection of uncultured bacteria. *FEMS*
429 *Microbiol Lett* **229:** 23–30.

430 Koike S, Yabuki H & Kobayashi Y (2007) Validation and application of real-time
431 polymerase chain reaction assays for representative rumen bacteria. *Animal Science*
432 *Journal* **78,** 135–141.

433 Krause DO & Russell JB (1996) How many ruminal bacteria are there? *J Dairy Sci* **79:**
434 1467–1475.

435 Li M, Wang B, Zhang M *et al*., (2008) Symbiotic gut microbes modulate human metabolic
436 phenotypes. *Proc Natl Acad Sci* **105**:2117-2122.

437 Lilburn TG, Schmidt TM & Breznak JA (1999) Phylogenetic diversity of termite gut
438 spirochaetes. *Environ Microbiol* **1:**331–345.

439 Ozutsumi Y, Tajima K, Takenaka A & Itabashi H (2005) The effect of protozoa on the
440 composition of rumen bacteria in cattle using 16S rRNA gene clone libraries. *Biosci*
441 *Biotechnol Biochem* **69**:499-506.

442 Paster BJ & Canale-Parola E (1985) *Treponema saccharophilum* sp. nov., a large pectinolytic
443 spirochete from the bovine rumen. *Appl Environ Microbiol* **50**:212- 219.

444 Paster BJ, Dewhirst FE, Weisburg WG, Fraser GJ, Tordo LA, Hespell RB, Stanton TB, Zablen
445 L & Woese CR (1991) Phylogenetic analysis of the spirochetes. *J Bacteriol* **173**, 6101-
446 6109.

447 Paster BJ, Dewhirst FE, Cooke SM, Fussing V, Poulsen LK & Breznak JA **(**1996) Phylogeny
448 of not-yet-cultured spirochetes from termite guts. *Appl Environ Microbiol* **62:**347–352.

449 Paster BJ, Boches SK, Galvin JL, Ericson RE, Lau CN, Levanos VA, Sahasrabudhe A &
450 Dewhirst FE (2001) Bacterial diversity in human subgingival plaque. *J Bacteriol*
451 **183:**3770–3783.

452 Piknova M, Wanda G, Miltko R, Peter J, Kasperowiscz A & Michalowski T (2008)
453 *Treponema zioleckii* sp. nov., a novel fructan utilizing species of rumen treponemes.
454 *FEMS Microbiol Lett* **289,** 166–172.
455 Saito N & Nei M (1987) The neighbour-joining method: a new method for constructing
456 phylogenetic trees. *Mol Biol Evol* **4**: 406–425.
457 Shannon CE & Weaver W (1949) The mathematical theory of communication. University of
458 Illinois Press. Urbana, IL.
459 Shinkai T, Ueki T & Kobayashi Y (2010) Detection and identification of rumen bacteria
460 constituting a fibrolytic consortium dominated by *Fibrobacter succinogenes*. *Animal*
461 *Science Journal* **81,** 72–79.
462 Stackebrandt E & Goebel EM (1994) Taxonomic note: a place for DNA-DNA reassociation
463 and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J*
464 *Syst Bactereriol* **44**:846-849.
465 Stanton TB & Canale-Parola E (1979) Enumeration and selective isolation of rumen
466 spirochetes. *Appl Environ Microbiol* **38**:965-973.
467 Stanton TB & Canale-Parola E (1980) *Treponema bryantii* sp. nov., a rumen spirochete that
468 interacts with cellulolytic bacteria. *Arch Microbiol* **127**:145-156.
469 Stevenson DM & Weimer PJ (2007) Dominance of *Prevotella* and low abundance of
470 classical ruminal bacterial species in the bovine rumen revealed by relative quantification
471 real-time PCR. *Appl Microbiol Biotechnol* **75**: 165–174.
472 Sundset MA, Præsteng KE, Cann IK, Mathiesen SD & Mackie RI (2007) Novel rumen
473 bacterial diversity in two geographically separated sub-species of reindeer. *Microbial*
474 *Ecol* **54:** 424-438.
475 Tajima K, Aminov RI, Nagamine T, Ogata K, Nakamura M & Benno Y (1999) Rumen
476 bacterial diversity as determined by sequence analysis of 16S rDNA libraries. *FEMS*
477 *Microbiol Ecol* **29**: 159–169.

478 Tajima K, Aminov RI, Nagamine T, Matsui H, Nakamura M & Benno Y (2001) Diet-
479 dependent shifts in the bacterial population of the rumen revealed with real-time PCR.
480 *Appl Environ Microbiol* **67**: 2766–2774.
481 Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: molecular evolutionary genetics
482 analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**:1596-1599.
483 Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DG **(**1997) The
484 CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment
485 aided by quality analysis tools. *Nucleic Acid Res* **25**:4876–4882
486 Watanabe K, Kodama Y & Harayama S (2001) Design and evaluation of PCR primers to
487 amplify bacterial 16S ribosomal DNA fragments used for community fingerprinting. *J*
488 *Microbiol Methods* **44**:253-62.
489 Welkie DG, Stevenson DM & Weimer PJ (2010) ARISA analysis of ruminal bacterial
490 community dynamics in lactating dairy cows during the feeding cycle. *Anaerobe* **16**: 94-
491 100.
492 Whitford MF, Forster RJ, Beard CE, Gong JH & Teather RM (1998) Phylogenetic analysis of
493 rumen bacteria by comparative sequence analysis of cloned 16S rRNA genes. *Anaerobe*
494 **4**: 153–163.
495 Wojciechowicz M & Ziolecki A (1979) Pectinolytic enzymes of large rumen treponemes.
496 *Appl Environ Microbiol* **37**:136-142.
497 Yang S, Ma S, Chen J, Mao H, He Y, Xi D, Yang L, He T & Deng W (2010) Bacterial
498 diversity in the rumen of Gayals (Bos frontalis), Swamp buffaloes (Bubalus bubalis) and
499 Holstein cow as revealed by cloned 16S rRNA gene sequences. *Mol Biol Rep* **37**:2063–
500 2073.
501 Yu Y, Breitbart B, McNairnie P & Rohwer F (2006) FastGroupII: A web-based
502 bioinformatics platform for analyses of large 16S rDNA libraries. *BMC Bioinformatics*
503 **7**:57.

504 Yu Z & Morrison **(**2004) Improved extraction of PCR-quality community DNA from digesta
505 and fecal samples. *BioTechniques* **36:**808-812.
506 Ziolecki A (1979) Isolation and characterization of large treponemes from the bovine rumen.
507 *Appl Environ Microbiol* **37**:131-135.
508 Ziolecki A & Wojciechowicz M (1980) Small pectinolytic spirochetes from the rumen. *Appl*
509 *Environ Microbiol* **39**:919-922.
510
511

Table 1 PCR primers used in this study

Target	Primer sequences (5'-3')		Annealing Temperature (^{o}C)	Extension time (sec)	Product Size (bp)	Efficiency	Reference
Total bacteria	BAC341F	ACTCCTACGGGAGGCAG	57	19	465	1.92	[a]
	BAC 805R	GACTACCAGGGTATCTAATCC					[a]
Rumen *Treponema*-group	g-TrepoF	GGCAGCAGCTAAGAATATTCC	64	23	575	1.91	This study
	BAC926R	CCGTCAATTCCTTTGAGTTT					[b]
Treponema bryantii	T. bryF	AGTCGAGCGGTAAGATTG	57	18	421	1.95	[c]
	T. bryR	CAAAGCGTTTCTCTCACT					[c]

[a], Stevenson and Weimer (2007); [b], Watanabe *et al.* (2001); [c], Tajima *et al.* (2001)

Table 2 Percentages of *Treponema* relative to total bacteria in the rumen of sheep fed a hay or concentrate diet as determined by real-time PCR

	Alfalfa	Orchardgrass	Concentrate	
Target	Mean±SD	Mean±SD	Mean±SD	*P* value
Rumen *Treponema*-group	1.05± 0.62	0.57 ± 0.26	0.92± 0.85	0.648
Treponema bryantii	0.02± 0.01	0.02 ± 0.01	0.02 ± 0.03	0.977

561 **Table 3** Sequence diversity and coverage estimation of individual and combined libraries

	Alfalfa	Orchardgrass	Concentrate	Combined
Total No. of clones	111	101	101	313
Total No. of OTUs	32	32	30	67
Chao 1	68.6	41.7	60.9	101
Shannon index	3.22	3.12	2.88	3.64
Good's coverage (%)	83.8	88.1	85.2	91.1
Cultured *Treponema*				77 (2 OTUs)
T. bryantii, ∫(OTU)	47 (1 OTU)	19 (1 OTU)	10 (1 OTU)	76 (1 OTU)
T. saccharophilum, (OTU)	0	0	0	0
T. zioleckii, (OTU)	1 (1OTU)	0	0	1 (1 OTUs)
Uncultured *Treponema**				236 (65 OTUs)
Previously reported, (OTU)	20 (8 OTUs)	22 (6 OTUs)	41 (6 OTUs)	70 (11 OTUs)
Not reported, (OTU)	43 (22 OTUs)	60 (25 OTUs)	50 (23 OTUs)	166 (54 OTUs)

562 *Clones showing 86-96% sequence similarity with any sequence in the NCBI database were considered as
563 Not reported, while those having 97% or more similarity with uncultured clones were grouped as previously reported.
564 ∫Numbers in parentheses indicate predicted OTUs from the clones.
565
566
567
568
569
570
571
572
573
574
575
576
577

578 **Figure legends**

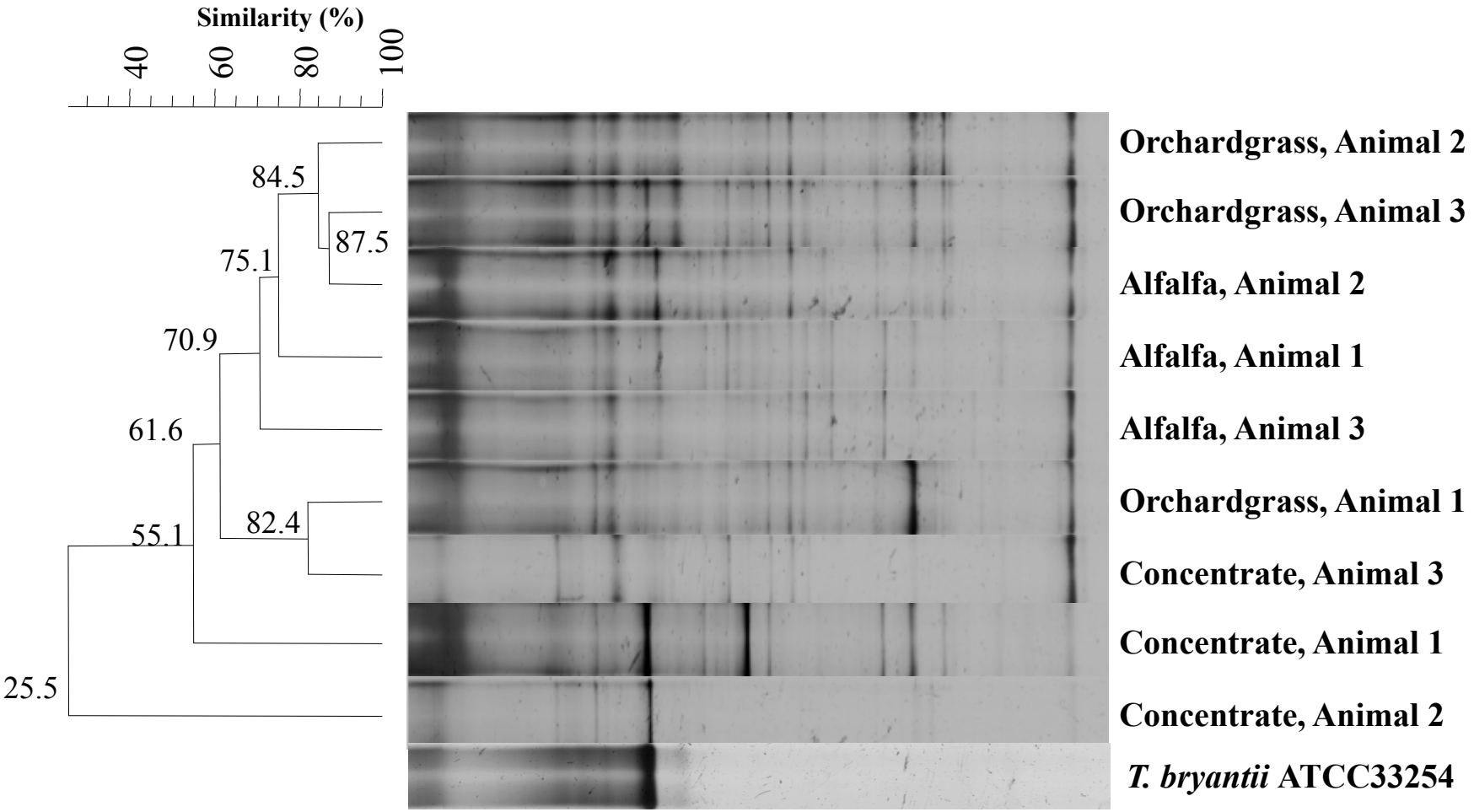
579 **Fig. 1** DGGE profiles of rumen *Treponema* 16S rRNA genes derived from rumen samples of
580 sheep fed different diets. Dendrograms were constructed using the unweighted pair group
581 method with arithmetic mean clustering analysis. The figure includes DGGE fingerprints of *T.*
582 *bryantii* ATCC33254. Numbers (1 to 3) indicate animal ID.

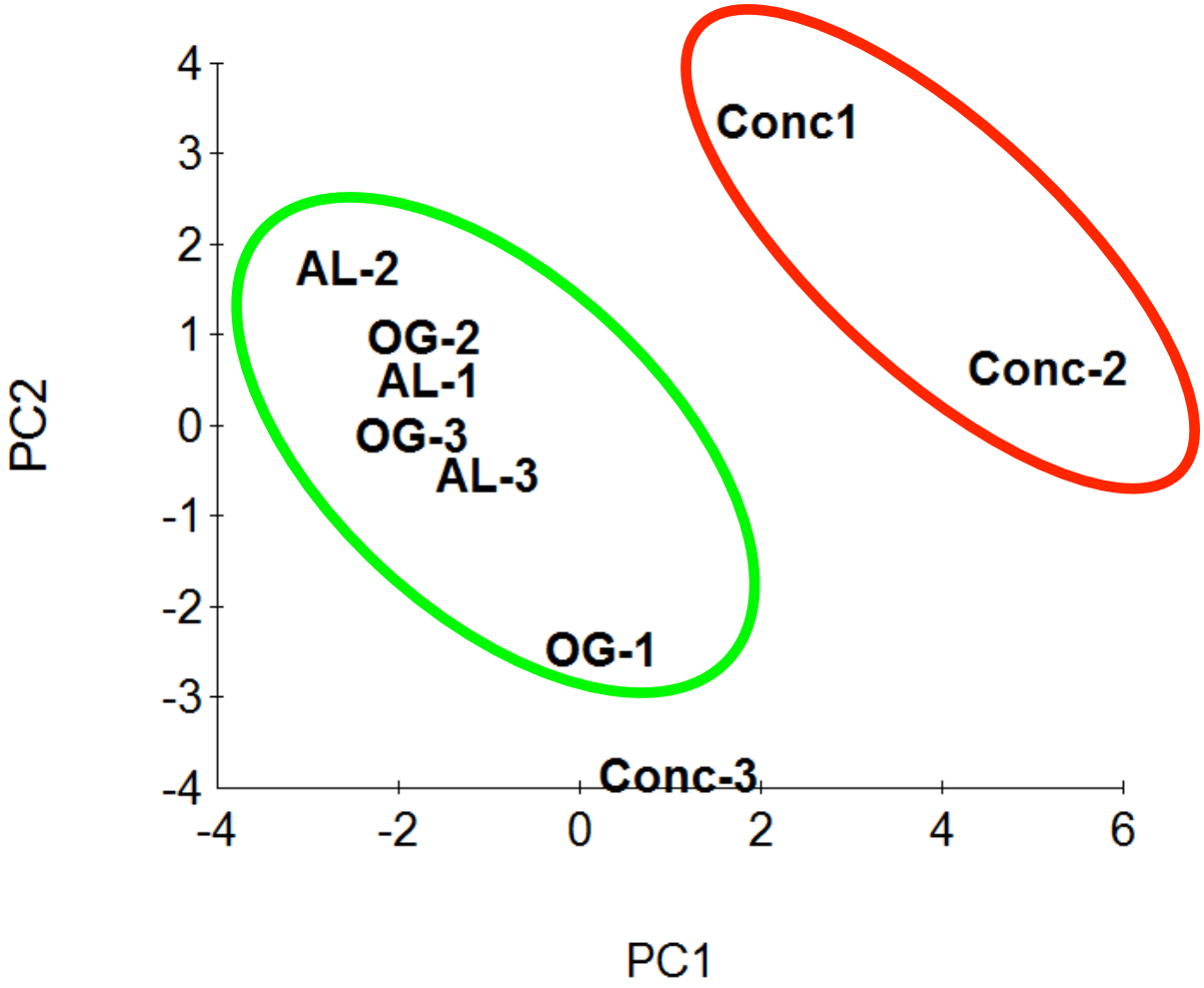
583

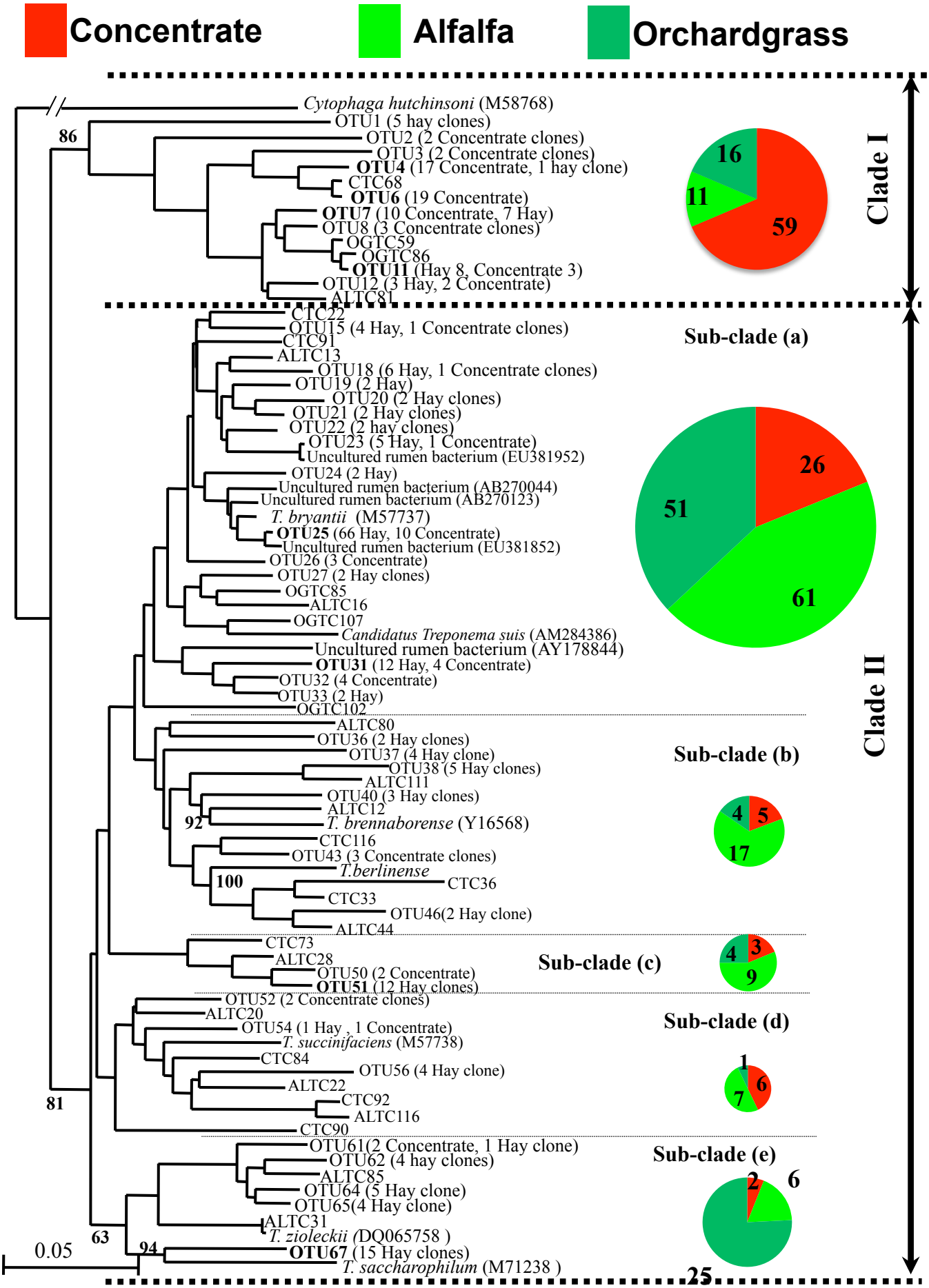
584 **Fig. 2** Principal component analysis (PCA) of DGGE fingerprints showing separate clusters
585 for the hay (AL-Alfalfa, OG-Orchardgrass) and concentrate (Conc) diets. Numbers (1 to 3)
586 indicate animal ID.

587

588 **Fig. 3** Phylogenetic placement of rumen *Treponema* 16S rRNA gene clone sequences
589 retrieved from sheep fed alfalfa, orchardgrass, and concentrate based diets. For the alfalfa,
590 orchardgrass, and concentrate diets, clone names respectively begin with ALTC, OGTC, and
591 CTC followed by clone number. The tree includes sequences of cultured rumen *Treponema*
592 species and other uncultured rumen *Treponema* clones. Clones having < 97% sequence
593 similarity were considered to belong to a distinct OTU; the number of clones in each OTU is
594 indicated in brackets. OTUs containing more than 10 clones are shown in bold. The pie charts
595 in each clade show the distribution of clones from each library. The size of the pie chart
596 corresponds with the number of clones in the clade. Bootstrap values above 50% are shown
597 as the percentage of 1000 replicates. The horizontal bar represents nucleotide substitutions
598 per sequence position.


Similarity (%)
40
60
80
100
84.5
87.5
75.1
70.9
61.6
82.4
55.1
25.5
Orchardgrass, Animal 2
Orchardgrass, Animal 3
Alfalfa, Animal 2
Alfalfa, Animal 1
Alfalfa, Animal 3
Orchardgrass, Animal 1
Concentrate, Animal 3
Concentrate, Animal 1
Concentrate, Animal 2
T. bryantii ATCC33254


Conc1
Conc-2
AL-2
OG-2
AL-1
OG-3
AL-3
OG-1
Conc-3
PC2
PC1
4
3
2
1
0
-1
-2
-3
-4
-4
-2
0
2
4
6


Concentrate
Alfalfa
Orchardgrass
Cytophaga hutchinsoni (M58768)
OTU1 (5 hay clones)
OTU2 (2 Concentrate clones)
OTU3 (2 Concentrate clones)
OTU4 (17 Concentrate, 1 hay clone)
CTC68
OTU6 (19 Concentrate)
OTU7 (10 Concentrate, 7 Hay)
OTU8 (3 Concentrate clones)
OGTC59
OGTC86
OTU11 (Hay 8, Concentrate 3)
OTU12 (3 Hay, 2 Concentrate)
ALTC81
86
16
11
59
Clade I
CTC22
OTU15 (4 Hay, 1 Concentrate clones)
CTC91
ALTC13
OTU18 (6 Hay, 1 Concentrate clones)
OTU19 (2 Hay)
OTU20 (2 Hay clones)
OTU21 (2 Hay clones)
OTU22 (2 hay clones)
OTU23 (5 Hay, 1 Concentrate)
Uncultured rumen bacterium (EU381952)
OTU24 (2 Hay)
Uncultured rumen bacterium (AB270044)
Uncultured rumen bacterium (AB270123)
T. bryantii (M57737)
OTU25 (66 Hay, 10 Concentrate)
Uncultured rumen bacterium (EU381852)
OTU26 (3 Concentrate)
OTU27 (2 Hay clones)
OGTC85
ALTC16
OGTC107
Candidatus Treponema suis (AM284386)
Uncultured rumen bacterium (AY178844)
OTU31 (12 Hay, 4 Concentrate)
OTU32 (4 Concentrate)
OTU33 (2 Hay)
OGTC102
Sub-clade (a)
26
51
61
ALTC80
OTU36 (2 Hay clones)
OTU37 (4 Hay clone)
OTU38 (5 Hay clones)
ALTC111
OTU40 (3 Hay clones)
ALTC12
92
T. brennaborense (Y16568)
CTC116
OTU43 (3 Concentrate clones)
T.berlinense
100
CTC36
CTC33
OTU46(2 Hay clone)
ALTC44
Sub-clade (b)
4
5
17
CTC73
ALTC28
OTU50 (2 Concentrate)
OTU51 (12 Hay clones)
Sub-clade (c)
4
3
9
OTU52 (2 Concentrate clones)
ALTC20
OTU54 (1 Hay , 1 Concentrate)
T. succinifaciens (M57738)
CTC84
OTU56 (4 Hay clone)
ALTC22
CTC92
ALTC116
81
CTC90
Sub-clade (d)
1
6
7
OTU61(2 Concentrate, 1 Hay clone)
OTU62 (4 hay clones)
ALTC85
OTU64 (5 Hay clone)
OTU65(4 Hay clone)
ALTC31
T. zioleckii (DQ065758)
63
94
OTU67 (15 Hay clones)
T. saccharophilum (M71238)
0.05
Sub-clade (e)
2
6
25
Clade II