Diversity of formyltetrahydrofolate synthetase genes in the rumens of roe deer (Capreolus pygargus) and sika deer (Cervus nippon) fed different diets
Diversity of formyltetrahydrofolate synthetase genes in the rumens of roe deer (Capreolus pygargus) and sika deer (Cervus nippon) fed different diets

Zhipeng Li\textsuperscript{a,}\#, Gemma Henderson\textsuperscript{b}, Yahan Yang\textsuperscript{a}, Guangyu Li\textsuperscript{a,}\#

a. Jilin Provincial Key Laboratory for Molecular Biology of Special Economic Animals, Institute of Special Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, Changchun, Jilin, China

b. AgResearch, Grasslands Research Centre, Palmerston North 4442, New Zealand

\textbf{# Corresponding authors}

Zhipeng Li, Email: zhplicaas@163.com, Tel: (+86)-43181919515, Fax: (+86)-43181919800.

No 4899, Juyue street, Changchun, Jilin Province, 130112, China.

Guangyu Li, Email: tcslgy@126.com, Tel: (+86)-43181919805. Fax: (+86)-43181919800.

No 4899, Juyue street, Changchun, Jilin Province, 130112, China.

\textbf{Email addresses:}

Zhipeng Li: zhplicaas@163.com,

Gemma Henderson: gh.microbiology@gmail.com

Yahan Yang: yangyahan@caas.cn

Guangyu Li: tcslgy@126.com
Abstract

Reductive acetogenesis by homoacetogens represents an alternative pathway to methanogenesis to remove metabolic hydrogen during rumen fermentation. In this study, we investigated the occurrence of homoacetogen in the rumens of pasture-fed roe deer (*Capreolus pygargus*) and sika deer (*Cervus nippon*) fed either oak leaf (tannin-rich, 100 mg/kg dried matter), corn stover, or corn silage-based diets, by using formyltetrahydrofolate synthetase (FTHFS) gene sequences as a marker. The diversity and richness of FTHFS sequences was lowest in animals fed oak leaf, indicating that tannin-containing plants may affect rumen homoacetogen diversity. FTHFS amino acid sequences in the rumen of roe deer significantly differed from those of sika deer. The phylogenetic analyses showed that 44.8% of sequences in pasture-fed roe deer, and 72.1%, 81.1%, and 37.5% of sequences in sika deer fed oak leaf, corn stover, and corn silage based diets, respectively, may represent novel bacteria that have not yet been cultured. These results demonstrate that the rumens of roe deer and sika deer harbor potentially novel homoacetogens and that diet may influence homoacetogen community structure.

**Keywords:** Homoacetogens; Deer; Diets
Introduction

Enteric methane emissions from livestock are a significant source of global agricultural greenhouse gases, accounting for 2.1 GtCO$_2$ Eq/year (6.3%) (Opio 2013; Smith P. and C. Mbow 2014). Methane emitted from ruminants also represents a loss of between 2 and 12% of the ingested feed energy (Johnson and Johnson 1995). In rumen fermentation, feed components are first digested by numerous enzymes leading to the release of monomers, which are then fermented by the rumen microbiota, resulting in large amounts of metabolic hydrogen being produced (Hill et al. 2016).

The removal of metabolic hydrogen facilitates the reduction and reoxidation of essential enzyme cofactors, such as NADH to NAD$^+$. If hydrogen accumulates, the rate of fermentation, and in turn animal productivity, can decrease (Thauer et al. 2010). Hydrogen is used in methanogenesis by rumen archaea (also known as methanogens), a process in which methane is formed (Hook et al. 2010). Re-directing metabolic hydrogen away from methanogenesis towards other pathways is important to reduce greenhouse gas emissions and improve feed efficiency.

Reductive homoacetogenesis also represents a process by which hydrogen can be removed. It can lead to a 0-26% improvement in energy harvesting through the production of acetate (Gagen et al. 2012), especially when methanogens have not been established or are inhibited in the rumen (Fonty et al. 2007). Homoacetogens have been reported to occur in the gastrointestinal tract of herbivorous animals. Many of these homoacetogens have yet to be characterized and may play a role in hydrogen disposal (Denman et al. 2015; Gagen et al. 2010; Gagen et al. 2012; Gagen et al. 2014;
Homoacetogens isolated from the fore-stomach of low-methane-emitting kangaroos suggest that homoacetogenesis may play a major role in hydrogen removal (Godwin et al. 2014; Ouwerkerk et al. 2009). Roe deer (*Capreolus pygargus*) emit less methane than other ruminants (Crutzen et al. 1986). It is known that host genetics significantly affects rumen ecology (Russell and Rychlik 2001). Therefore, these results suggest that there may be novel homoacetogens in the rumen of roe deer, however, to date, these have not been characterised.

Diet is another significant factor affecting rumen microbiota composition (Henderson et al. 2015). Tannins are polyphenolic compounds, which are widely distributed in the diets of herbivores (McSweeney et al. 2001). Previous studies have shown that tannins not only reduced methane production in ruminants (Jayanegara et al. 2012; Tan et al. 2011), but also altered the diversity and composition of rumen methanogens (Min et al. 2014). Our previous study also demonstrated that tannin-containing plants significantly affected the rumen bacterial and methanogen communities and fermentation patterns in the rumens of sika deer (Li et al. 2015; Li et al. 2013). In considering the important roles of the rumen bacteria and methanogens in hydrogen metabolism, tannin-rich diets may affect the diversity and community structure of homoacetogens.

The aim of this study, was to use the formyltetrahydrofolate synthetase (FTHFS) gene (*fhs*), which encodes a structurally and functionally conserved enzyme in the acetyl-CoA pathway (Lovell and Leaphart 2005), as a functional marker to study the
diversity of homoacetogens in the rumens of roe deer and sika deer; and to examine the effect of host and diet on homoacetogen community structure.

**Materials and Methods**

Animals and sampling

The use of roe deer (*Capreolus pygargus*) and sika deer (*Cervus nippon*) in this study was approved and authorized by the Chinese Academy of Agricultural Sciences Animal Care and Use Committee and by the Institute of Special Animal and Plant Sciences Wild Animal and Plant Subcommittee. These animals were also used in our previous studies (Li et al. 2015; Li et al. 2014).

Three rumen-cannulated adult male sika deer, maintained at the research farm of the Institute of Special Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, in Jilin Province, were used in a 3×3 Latin square design. A concentrate diet (64.5% corn, 19.7% soybean meal, 12.8% distiller dried grains with solubles and a 3% mixture of vitamins and mineral salts), that was supplemented with either oak leaf (*Xylosmaracemosum*; tannin content, 100 mg/kg dried matter), corn stover, or corn silage, formed three different diets. The ratio of forage to concentrate was 50:50 on the basis of dry matter. Each sika deer was fed oak leaf, corn stover, or corn silage-based diets in turn (*Table S1*). All Sika deer were fed twice each day at 8:00 AM and 4:00 PM and had free access to water. Following a one-week adaption to the diets, and after receiving each diet for 28 days, rumen contents were obtained before morning feeding. By the end of the study, we had collected nine samples for all diets (3 samples per diet).
Three free-range, healthy, male roe deer, reared by a local farmer on grazing pasture and maintained in local mountains of Chifeng City, Inner Mongolia Autonomous Region in China, were also used in this study. The animals were slaughtered before morning feeding and liquid and solid rumen contents obtained. After sampling, rumen samples were immediately frozen in liquid nitrogen and stored at -80°C until further analysis.

**DNA extraction, FTHFS amplification and clone libraries**

Total genomic DNA was extracted from the whole ruminal contents of each animal using a QIAamp DNA Stool Mini Kit (QIAGEN, Valencia, CA) according to the manufacturer’s instructions. Partial FTHFS gene sequences were amplified using the primers described by Leaphart and Lovell (Leaphart and Lovell 2001). Briefly, each 25μl reaction contained 50 ng template DNA, 0.25 mM of each primer, 250 mM dNTPs, 0.5 U of Ex Taq and 2.5μl Ex Taq buffer (TaKaRa, Dalian, China). A touchdown PCR was performed on a EasyCycler 96 (Analytik Jena AG, Jena, Germany) according to the protocol of Leaphart & Lovell (2001). The touchdown protocol consisted of a initial denaturation at 94°C for 2 min, followed by 9 cycles of 94°C for 30 s, 63°C for 30 s (decreased by 1°C per cycle to 55°C), and 72°C for 30 s, followed by 25 cycles at an annealing temperature of 55°C, and a final extension of 72°C for 2 min. PCR products were assessed using 2% agarose gel electrophoresis (approximately 1.1 kb), and were purified using a QIAquick PCR Purification Kit (QIAGEN, Valencia, CA). The concentrations of PCR products of each animal from each group were measured by spectrophotometry (SPECORD 50, Analytik Jena,
Germany). Equal amounts of PCR-amplified products from each animal in each group were pooled together, and were then used to construct clone libraries. Four FTHFS gene clone libraries were constructed from the pooled PCR products using the TOPO® TA Cloning® Kit (Invitrogen, Carlsbad, CA). Positive clones were sequenced using an ABI 3730XL DNA Analyzer.

Analysis of FTHFS sequences

FTHFS sequences were translated and aligned with ClustalW within the MEGA 5.0 software package (Tamura et al. 2011). The alignment and translation were corrected manually based on previously identified conserved or functionally important areas of FTHFS (Henderson et al. 2010). The chimera check program Bellerophon was used to identify chimeric sequences (Huber et al. 2004). FTHFS amino acid sequences were clustered into operational taxonomic units (OTUs) at a distance of ≤ 0.025 using MOTHUR (Gagen et al. 2010; Schloss et al. 2009). Although OTUs for FTHFS amino acid sequences do not correspond to a species, 97.5% identity was chosen in order to group similar FTHFS sequences (Gagen et al. 2010). Libshuff analysis was performed to compare the structure of fhs libraries (deduced amino acids) between the four groups based on evolutionary distance of all sequences. A series of likelihood ratio tests were performed using ProtEST (Cuff et al. 2000) to determine the best substitution model (LG+G) for amino acid sequences. Maximum likelihood phylogenies were generated using the LG+G model in RAxML v8.0.0 with 1,000 bootstrap replicates (Stamatakis 2014).

Homoacetogen similarity (HS) scores for deduced FTHFS sequences were
calculated by examining the conserved amino acid residues in FTHFS from acetogens using the method described previously (Henderson et al. 2010). Additionally, a homoacetogen FTHFS profile hidden Markov model (HoF-HMM) was used to see whether FTHFS sequences were similar to those of homoacetogens, by calculating HMMER bit scores of sequences as described previously (Henderson et al. 2010).

*Nucleotide sequence accession numbers*

All nucleic acid sequences obtained in this study were deposited in GenBank under the accession numbers KP144455-KP144781.

**Results**

*Comparison of FTHFS sequences between different animals and diets*

Of 350 total sequences, 23 were identified as chimeric and excluded from further analyses. The remaining 327 non-chimeric sequences were classified into 54 operational taxonomic units (OTUs with 98% sequence similarity). The Good's coverage of FTHFS sequence clone library, which was measured using the mothur platform, was greater than 80%. Libshuff analyses show that FTHFS amino acid sequences in the rumen of roe deer significantly differed from those in the rumen of sika deer fed three forage-based diets ($\Delta$CXY score ≤ 0.02, significance ≤ 0.0001). The differences between oak leaf and corn stover ($\Delta$CXY score ≤ 0.001, significance ≤ 0.01), between oak leaf and corn silage were significant ($\Delta$CXY score ≤ 0.02, significance ≤ 0.003), but, the difference between corn stover and corn silage was not significant ($\Delta$CXY score ≤ 0.02, significance ≤ 0.9). Interestingly, only two of the 54 OTUs were shared by all four groups (Figure 1). We also found that the diversity and
richness were lowest in sika deer fed oak leaf-based diets (Table S2).

Phylogenetic analysis of putative FTHFS sequences

FTHFS sequences broadly formed five clusters (I, II, III, IV, and V), which could be further grouped into 19 sub-clusters (Figure 2). Of the sika deer fed corn stover and corn silage, 3.9% and 3.1% of FTHFS sequences grouped in Cluster I and were most similar to FTHFS of *Clostridium acetobutylicum* (amino acid identity: 59%-60%). Sequences of roe deer and oak leaf groups were not found in this cluster.

Within cluster II, rumen sub-clusters 5 and 23 represented 3.4%, 3.9% and 5.2% of the total FTHFS amino acid sequences in roe deer, corn stover, and corn silage groups, respectively. Sequences in these two sub-clusters grouped with FTHFS sequences of the homoacetogens *Clostridium magnum*, *Clostridium aceticum*, *Clostridium formicaceticum*, and *Thermoanaerobacter kivui* as supported by high bootstrap values (amino acid identity: 69%-79%).

Within cluster III, there were 14 sub-clusters and two sequences (SdCI 71 and SdCS 78) that did not group in a sub-cluster, representing 16.1%, 13.9%, 20.8% and 21.9% of the total FTHFS amino acid sequences in the roe deer, oak leaf, corn stover, and corn silage groups, respectively (Figure 2). These sequences were most closely related to FTHFS sequences of *Treponema* spp. and *Blautia* spp. (amino acid identity: 69%-76%). Within cluster IV, rumen cluster 6 represented 25.3%, 58.2%, 56.4% and 10.4% of all FTHFS sequences in roe deer, oak leaf, corn stover, and corn silage groups, respectively. These sequences grouped nearest to *Sporomusa termitida* and *Sporomusa ovata* (amino acid identity: 70%-72%).
The remaining sequences grouped in cluster V (roe deer: 55.2%; oak leaf: 27.9%; corn stover: 14.9%; corn silage: 59.4%) in the lower half of FTHFS tree, could be further classified into 13 sub-clusters, which were related to sulfate-reducing and purinolytic bacteria, including *Clostridium acidurici*, *Clostridium cylindrosporum*, *Firmicutes* spp. and *Butyrivibrio* spp. (amino acid identity: 57%-61%).

**Similarity of FTHFS sequences to those of known homoacetogens**

Functional analyses of FTHFS sequences including the homoacetogen similarity (HS) scores and FTHFS profile hidden Markov model (HoF-HMM) (Henderson et al. 2010), were used to characterize sequences of uncultured bacteria and compare them with those of known homoacetogens (*Figures 2 and 3*). The results of HS score and HoF-HMM can be used to determine whether sequences likely originated from true homoacetogens.

FTHFS sequences that fell into clusters I had low HS scores (37.5%~70%), and FTHFS sequences within cluster V had HS scores in the range of 81.3% to 81.7%. Moreover, these sequences had intermediate HoF-HMM bit scores (~ 650, *Figure 3*). FTHFS amino acid sequences within cluster II displayed high HS scores (83.8%-95%) and had the best overall HoF-HMM bit scores, ranging from 696 to 764.9. FTHFS sequences within clusters III and IV did not group with FTHFS sequences of known homoacetogens, but had relatively high HoF-HMM bit scores, ranging from 656.4 to 772.4 and HS scores in the range of 82.5%-97.5%.

**Discussion**

In this study, we examined homoacetogen communities in the rumens of
pasture-fed roe deer, and sika deer fed oak leaf, corn stover and corn silage based diets using fhs gene sequences, which encode formyltetrahydrofolate synthetase.

The FTHFS sequence coverage indicated that the majority of the FTHFS sequences in the rumen of roe deer and sika deer were captured and that there may be more FTHFS sequences at an approximate species level to be uncovered. Interestingly, we found that the FTHFS sequences diversity and composition in the rumen of sika deer fed oak leaf-based diets significantly differed from those of other groups (Figure 1 and Table S2). Mitsumori et al. (2014) showed that the diversity and structure of homoacetogens in the rumen of cattle was altered when the methane production was decreased by bromochloromethane. These results indicate that the tannin-containing plants also affected the homoacetogen community. In addition, previous studies demonstrated that tannins inhibited many rumen fibrolytic and proteolytic bacteria through binding of their cell membranes, and also reduced the numbers of protozoa, which in turn affected hydrogen production and can lead to decreased methane emissions (Bae et al. 1993; Jayanegara et al. 2012; Puchala et al. 2005; Smith et al. 2005; Tan et al. 2011).

Phylogenetic analysis was used to characterize homoacetogen populations in the rumens of roe deer and sika deer fed different diets (Figure 2). FTHFS sequences within clusters I and V were phylogenetically related to sulfate-reducing and purinoloytic bacteria, including Clostridium acetobutylicum, Clostridium acidurici, and Clostridium cylindrosporum. Previous studies reported that FTHFS can also be used as a methyltransferase in purine and glycine degradation and in the metabolism of
some sulfate-reducing bacteria (Fuchs 1986). *Clostridium acidurici* and *C. cylindrosporum* are purine-degrading microorganisms that play a role in transfer of a formimino group to tetrahydrofolate (THF), which is further converted to formyl-THF (Barker and Beck 1942). Finally, FTHFS contributes to the generation of ATP together with the release of formate and THF (Barker and Beck 1942; Ottesen and Leadbetter 2010). Therefore, this may be related to the findings of previous study showing that FTHFS has different substrate specificities (Ottesen and Leadbetter 2011). It has been reported that the primers used in this study (Leaphart and Lovell 2001) sometimes do not detect *Blautia* sp. Ser8, *Blautia hydrogenotrophica*, *Blautia schinkii*, *Acetitomaculum ruminis*, *Oxobacter pfennigii*, and *Syntrophococcus sucromutans* (Gagen et al. 2010; Henderson et al. 2010). These results suggest the limitation of the choice of FTHFS primers alone as marker gene in characterizing rumen homoacetogens, as they cannot be restricted to true homoacetogens, and involve the ambiguous classification between the homoacetogens and non-homoacetogens. Therefore, there may be unrecovered homoacetogens in the rumens of sika deer and roe deer. Moreover, these findings suggest the primer acsB, which is unique in the acetyl-CoA pathway (Gagen et al., 2010), could also be applied to examine homoacetogen communities in future studies.

FTHFS sequences within cluster II could be derived from homoacetogenic bacteria as they grouped closely with FTHFS sequences of known homoacetogens (Drake et al. 2013). FTHFS sequences within clusters III were phylogenetically close to FTHFS sequences of from *Treponema* spp., *Blautia* spp., *Sporomusa termitida* and
Sporomusa ovata (Figure 2). Similar sequences were detected in the rumens of a Holstein cow fed a mixture of pasture and soybean meal (Matsui et al. 2008), a Friesian-Jersey cross cow fed pasture (Henderson et al. 2010), steers fed pasture (Gagen et al. 2010), and the forestomachs of eastern and red kangaroos (Ouwerkerk et al. 2009), and tammar wallabies (Gagen et al. 2010). These results suggest that these uncultivated groups are widespread in the gastrointestinal tract of herbivorous animals, regardless of diet or location. Typically different types of FTHFS sequences predominate in the guts of termites, cockroaches, and giant pandas (Ottesen and Leadbetter 2010, 2011; Pester and Brune 2006; Tune et al. 2014), suggesting that the evolutionary lifestyle of hosts maybe a factor that affects homoacetogen communities. This hypothesis warrants further investigation using widely distributed herbivorous animals, in particular ruminants with differing methane emissions.

FTHFS sequences in cluster III occurred less in the rumens of sika deer fed oak leaf-based diets (13.9%) than they did in the rumens of sika deer fed corn stover (19.8%) and corn silage (22.9%) based diets (Figure 2). Our previous results demonstrated that unclassified bacteria in the family Ruminococcaceae were higher in the rumens of sika deer fed corn stover and corn silage based diets, while unclassified bacteria belonging to the family Succinivibrionaceae were abundant in the rumens of sika deer fed oak leaf based diets (Li et al. 2015). Ruminococcaceae are presumed to play key roles in the degradation of plant material, producing hydrogen during this process (Rychlik and May 2000). Some Succinivibrionaceae produce succinate as the principal fermentation product, and some exogenous sources of hydrogen could
stimulate succinate formation (Lee et al. 1999; Pope et al. 2011; Stackebrandt and Hespell 2006). Taken together, these results indicate that hydrogen-producing microorganisms in the rumen may play a major role in shaping the homoacetogen community through the amounts of metabolic hydrogen available. However, this hypothesis needs to be further studied.

The low HS and intermediate HoF-HMM bit scores (~ 650) of FTHFS sequences in clusters I and V suggests that these sequences may not originate from homoacetogens (Figure 3). FTHFS sequences within clusters III and IV displayed considerably higher HoF-HMM bit scores and HS scores. Recent studies have shown that the HS scores of homoacetogens Clostridium sp. M62/1 clade and Thermacetogenium phaeum were between 75% and 95% (Ottesen and Leadbetter 2010) and 79% and 100% (Hori et al. 2011), respectively.

The present study is the first to examine homoacetogen community structures in the rumens of roe deer and sika deer fed different forage-based diets. The results suggest that novel homoacetogens are present in the rumens of sika and roe deer and that tannin-containing plants decrease the diversity and alter the composition of FTHFS sequences.

Acknowledgements

This study was supported by Natural Science Foundation of China (NSFC) (Grant No. 31501984). We especially thank the farmers at the research farm of the Institute of Special Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, for their assistance with collecting samples.
Conflict of Interest

No conflict of interest has been declared.
References


Henderson, G., Cox, F., Ganesh, S., Jonker, A., Young, W., and Janssen, P.H. 2015. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Scientific Reports 5: 14567.


Environ. Microbiol. 64(9): 3429-3436.


Figure legends

**Figure 1.** Venn diagram showing shared and unique formyltetrahydrofolate synthetase (FTHFS) sequence operational taxonomic units (OTUs) in the rumens of pasture-fed roe deer and sika deer fed oak leaf, corn stover and corn silage-based diets.

**Figure 2.** Phylogenetic tree of deduced formyltetrahydrofolate synthetase (FTHFS) amino acid sequences from the rumens of pasture-fed roe deer and sika deer fed oak leaf, corn stover and corn silage based diets. The tree was generated by maximum-likelihood distances between deduced amino acid sequences. Bootstrap values greater than 70% for both nucleotide and amino acid sequences are indicated at nodes. The scale bar represents 0.2 amino acid changes per alignment position. Clusters that contain known homoacetogenic, sulfate-reducing, and purinolytic bacteria and other cultured bacterial isolates are indicated by shading. The number of sequences (n) and the acetogens similarity (HS) score (expressed as a percentage) are indicated in parentheses as follows: (n, HS score).

**Figure 3.** Similarity of roe and sika deer rumen FTHFS sequences to a hidden Markov model constructed with the FTFHS sequences of known homoacetogens (HoF-HMM). The significance of the matches of deduced FTHFS sequences is expressed as an HMMER bit score; the higher the score, the better the match with HoF-HMM. A high HoF-HMM bit score could represent novel homoacetogens. Sequences obtained from the rumens of pasture-fed roe deer, and of sika deer fed oak leaf, corn stover and corn silage based diets are expressed as red, green, blue light and purple circles. The different HS score were represented by various shapes: HS score.
< 80% (black circle); 80% ≤ HS score < 90% (black triangle); HS score ≥ 90%
(black square). The higher HS score tended to have higher HMMER bit score.
Supporting Information

**Table S1** Study design for the 3×3 Latin square using three sika deer.

**Table S2** Coverage, diversity and richness indices in the roe deer, oak leaf, corn stover, and corn silage libraries.