Rumen microbial (meta)genomics and its application to ruminant production

Diego Morgavi, William Kelly, Peter Janssen, Graeme Attwood

Animal (2013), **7:s1**, pp 184–201 © The Animal Consortium 2012
doi:10.1017/S1751731112000419

Rumen microbial (meta)genomics and its application to ruminant production

D. P. Morgavi[1†], W. J. Kelly[2], P. H. Janssen[2] and G. T. Attwood[2]

[1]*INRA, UR1213 Herbivores, Site de Theix, F-63122 Saint-Genès-Champanelle, France;* [2]*Ruminant Nutrition and Animal Health, AgResearch, Private Bag 11008, Palmerston North 4442, New Zealand*

(Received 1 August 2011; Accepted 16 January 2012; First published online 1 March 2012)

Meat and milk produced by ruminants are important agricultural products and are major sources of protein for humans. Ruminant production is of considerable economic value and underpins food security in many regions of the world. However, the sector faces major challenges because of diminishing natural resources and ensuing increases in production costs, and also because of the increased awareness of the environmental impact of farming ruminants. The digestion of feed and the production of enteric methane are key functions that could be manipulated by having a thorough understanding of the rumen microbiome. Advances in DNA sequencing technologies and bioinformatics are transforming our understanding of complex microbial ecosystems, including the gastrointestinal tract of mammals. The application of these techniques to the rumen ecosystem has allowed the study of the microbial diversity under different dietary and production conditions. Furthermore, the sequencing of genomes from several cultured rumen bacterial and archaeal species is providing detailed information about their physiology. More recently, metagenomics, mainly aimed at understanding the enzymatic machinery involved in the degradation of plant structural polysaccharides, is starting to produce new insights by allowing access to the total community and sidestepping the limitations imposed by cultivation. These advances highlight the promise of these approaches for characterising the rumen microbial community structure and linking this with the functions of the rumen microbiota. Initial results using high-throughput culture-independent technologies have also shown that the rumen microbiome is far more complex and diverse than the human caecum. Therefore, cataloguing its genes will require a considerable sequencing and bioinformatic effort. Nevertheless, the construction of a rumen microbial gene catalogue through metagenomics and genomic sequencing of key populations is an attainable goal. A rumen microbial gene catalogue is necessary to understand the function of the microbiome and its interaction with the host animal and feeds, and it will provide a basis for integrative microbiome–host models and inform strategies promoting less-polluting, more robust and efficient ruminants.

Keywords: rumen, microbial, genomics, metagenomics, nutrition

Implications

Rumen microbes are essential for ruminant production. They allow ruminants to transform plant forages, inedible for humans, into high-quality foods. However, they are also responsible for methane production. Recent developments in the study of gut microbial communities (microbiomes) through genomics and metagenomics are revolutionising our understanding of the functions of the ecosystem and the interactions among their members and the host animal. A better knowledge of the rumen microbiome and its underlying functions through the construction of a gene catalogue can inform strategies to improve feed digestion efficiency and reduce enteric methane production, meeting the challenge of sustainability.

Introduction

The term 'superorganism' is used to describe the intimate association between a higher animal and its symbiotic microbiota (Lederberg, 2000; Goodacre, 2007). The gastrointestinal microbiota exerts protective, immunological, developmental and nutritional functions that benefit the host (Hooper, 2004) and can be justly considered an 'organ' with far more metabolic functions and capabilities than the host's tissues (Xu and Gordon, 2003; Egert *et al.*, 2006; Gill *et al.*, 2006). In humans, it has been shown that there are at least 3.3 million non-redundant microbial genes,

† E-mail: morgavi@clermont.inra.fr

a staggering number that is up to 150 times greater than the number of genes in the human genome (Qin *et al.*, 2010). In domestic herbivores, particularly in ruminants, the importance of symbiotic gastrointestinal microorganisms for nutrition, health and well-being of the host was recognised long before the equivalent benefit was apparent in man. Ruminants harvest energy from otherwise indigestible structural plant polysaccharides by providing a suitable environment for symbiotic gastrointestinal microbes. This is regulated so that only a partial, fermentative, degradation of the feed occurs, and the ruminant host can use the fermentation end products for its own nutrition. The host also profits from microbial metabolites such as vitamins and the high-quality proteins contained in microbes for its nutritional needs. From a microbiological point of view, the domestication of ruminants is a successful use of microorganisms by humans (Russell *et al.*, 2009). Farming of ruminants allowed earlier cultures and civilisations to have a stable food supply and to expand their living ranges to regions not suitable for crop production. These features are arguably attributed to ruminants' abilities to utilise the structural carbohydrates of plant fibres, making them both robust and adaptable to different climatic and feeding conditions. For instance, archaeological evidence from a neolithic farming site suggests that the number of domesticated ruminants rose, whereas that of pigs decreased, during a period of climate hardship that may have reduced the availability of other food resources (Balter, 2010).

The worldwide head count of domestic ruminants is 1.38×10^9 cattle, 1.96×10^9 sheep and goats and 2.21×10^8 buffaloes and camelids (FAO, 2009), which sustain directly or indirectly the livelihood of hundreds of millions of people worldwide. The economic value of ruminant products is important in both developing and developed countries. Ruminant meat and/or milk ranks at the top of the most important agricultural foods and commodities produced in every continent except in Asia, where the value of buffalo and cow milk is surpassed only by rice (FAO, 2009). Because of human population growth and increases in food consumption, the demand for meat and milk is expected to double in the next 40 years (FAO, 2006), inevitably boosting the number of ruminants on earth and posing a challenge to global resources and food security (World Bank, 2008). The rumen microbial symbionts give ruminants the capacity to feed on forages and other food not suitable for human consumption, therefore potentially allowing use of agricultural land not suitable for crops directly eaten by humans. However, the same ruminants are also responsible for producing the potent greenhouse gas methane and other pollutants such as nitrogen-rich wastes. Thus, the rumen microbiota epitomises both the good and the bad of ruminant production.

There is still a disconnection between rumen microbiology on the one hand and ruminant nutrition and production research on the other, with few studies addressing both aspects. It is true that knowledge of rumen microbes is not strictly necessary to feed and raise ruminants. However, a thorough understanding of the functions of the gastrointestinal microbiota and its interactions with the host animal is needed to produce meat and milk in a sustainable and environmentally sound way. The unprecedented wealth of information that can be generated using high-throughput 'omics' methods (second-generation genomic and metagenomic sequencing technologies, metatranscriptomics, metaproteomics and metametabolomics) to characterise the microbial component of the rumen ecosystem, combined with the sequenced genomes of cattle and sheep (The Bovine Genome Sequencing Analysis Consortium *et al.*, 2009; The International Sheep Genomics Consortium *et al.*, 2010), presents the opportunity to consider the ruminant superorganism as a whole system. This allows us to address, in a holistic way, important issues related to animal production such as the efficiency of forage digestion and ruminant methane emissions.

The purpose of this paper was to summarise recent advances in microbial genomics and metagenomics applied to the rumen, and to identify the necessary data that are still required to better comprehend its functioning and productivity.

The ruminant superorganism

The ruminant superorganism is a complex and interrelated system composed of many parts. The host and its microbiota are two major components that have co-evolved during millions of years, ensuring an increased fitness and increased chances of mutual survival. The composition of feed is another factor that profoundly influences the microbiota and host physiology (Ley *et al.*, 2008b; Clauss *et al.*, 2011). Herbivory is a successful strategy that has allowed mammals to expand the range of their habitat. In mammals, the diversification of modern herbivore lineages – the ancestors of our livestock – increased dramatically with the appearance of grasses $\sim$25 million years ago (Hume and Warner, 1980). As a result, herbivore species predominate among the total number of extant mammals and are among the largest terrestrial species (Stevens and Hume, 1998). Within herbivores, ruminants have enlarged their gastrointestinal tracts and increased the retention time of forages to facilitate the fermentation of feeds by their symbiotic microbes. Ruminants have evolved the most sophisticated system to harbour and take advantage of microbes in their forestomach (Stevens and Hume, 1998; Clauss *et al.*, 2010), making them very adaptable to a large variety of diets. This is a characteristic that was certainly an important domestication trait (Diamond, 1997 and 2002) and that can be attributed directly to microbes. The stratification of the rumen contents into gas, solid and liquid layers allows retention of particulate feeds for further processing while at the same time optimising the collection of microbial protein due to the high liquid turnover (Clauss *et al.*, 2010). Ruminants also have a strong innate immunity in the digestive surfaces of the remainder of the gastrointestinal tract, and have recruited and duplicated enzymes such as lysozymes and ribonucleases to play a digestive role, allowing a better utilisation of the microbial biomass leaving the rumen (Benner *et al.*, 2002; The Bovine Genome Sequencing Analysis Consortium *et al.*, 2009). These adaptive characteristics

evolved both to tolerate and make use of the products of this 'microbial organ'.

If the microbiome is considered as a single component from a nutritional perspective, the metabolic functions it can perform are bestowed by its different individual microbial members.

Microbes present in different anatomical parts of the gastrointestinal tract of animals were first described in 1843. The 'animalcules' observed by Gruby and Delafond (1843) with the techniques of their time were protozoa. Studies on gastrointestinal symbionts started in earnest 50 years ago with the development of anaerobic microbiology techniques by Robert E. Hungate, the father of modern anaerobic microbiology (Hungate, 1966). In addition to isolating and identifying microorganisms, Hungate also developed the concept of microbial ecology as his pioneering work on the rumen set the basis for the description and functional analysis of complex microbial ecosystems *in situ* (Konopka, 2006). Hungate's work is compiled in the classic text, *The Rumen and its Microbes*. The rumen is a complex ecosystem harbouring hundreds of phylotypes of bacteria, protozoa, fungi, methanogens and bacteriophages. These organisms interact with each other and with their environment, namely the host and feeds as stated above. Their concentration (up to 10^{11}, 10^{6}, 10^{6}, 10^{9} and 10^{10} cells or particles/ml for bacteria, protozoa, fungi, methanogens and bacteriophages, respectively) and hydrolytic activity is exceptionally high compared with terrestrial and aquatic ecosystems (Williams and Coleman, 1992; Hobson and Stewart, 1997; Mackie *et al*., 2002; Weimer *et al*., 2009). Similar to other ecosystems, the number of microbial species isolated and characterised from the rumen is low. It is estimated that less than 15% of rumen bacteria can be cultured using standard techniques (Teather, 2001; Edwards *et al*., 2004), highlighting the importance of molecular biology approaches to sidestep this limitation and study the rumen system *in toto*. The huge diversity of species in the rumen means that cultivation-based methods are simply not suited to follow changes in community structure, even if all the species could be cultured. The importance of molecular biology approaches becomes even more apparent when the intrinsic animal-to-animal variation is also considered. To make meaningful studies that uncover trends, large numbers of samples may need to be analysed at once.

Is there a need to completely characterise the microbiome? It is worth considering the diversity of the ecosystem and the difficulty of the challenge of understanding it. What will be the benefits to ruminant production of characterising the rumen microbiome? The basic functions of the rumen microbiome relevant to ruminant's nutrition and health, such as feed degradation, detoxification of plant toxins and feed contaminants, biotransformation of molecules of nutritional interest for humans or the production of methane, have been described (Hobson and Stewart, 1997). Likewise, some dietary practices, supplements and additives that modulate the rumen microbiota for increasing feed digestion and/or decreasing inefficient processes are recognised and are applied in some production systems (Nagaraja *et al*., 1997; Jouany and Morgavi, 2007; Chaucheyras-Durand *et al*., 2008; Martin *et al*., 2010). For instance, the rationale of using microbes with specific functions as probiotics (Ghorbani *et al*., 2002; Weiss *et al*., 2008) to improve the efficacy of rumen fermentation or the development of a vaccine against *Streptococcus bovis* (Gill *et al*., 2000) for reducing lactic acidosis are based on the knowledge gathered using traditional methods. However, despite the accumulated body of information, there is still an incomplete understanding of the functioning and ecology of the rumen microbiome and its behaviour cannot yet be accurately predicted. For instance, the complete mechanism of plant polysaccharide degradation, the quintessential rumen function, is not yet elucidated (Morrison *et al*., 2009; Wilson, 2011). Most of the work on fibre degradation in the rumen is based on three cultivable bacteria: *Fibrobacter succinogenes, Ruminococcus albus and Ruminococcus flavefaciens*. Because they are able to hydrolyse crystalline cellulose, they were considered to be important in the rumen (Flint *et al*., 2008; Wallace, 2008; Russell *et al*., 2009). The genomes of these three bacteria have been sequenced (see 'Rumen microbial genome projects section') and the enzymatic strategies to hydrolyse cellulose used by each species have been partially unravelled. Information on other rumen bacteria able to attack structural plant polysaccharides such as *Prevotella* and *Butyrivribrio* (Kelly *et al*., 2010; Purushe *et al*., 2010), as well as on some protozoa (Bera-Maillet *et al*., 2005) and the anaerobic fungi (Griffith *et al*., 2010), also exist. However, the microbial community performing the complex deconstruction of plant cell walls and how the components of the community interact cannot be fully described at the moment. *F. succinogenes*, *R. flavefaciens* and *R. albus* have been found to account for a mere 1% to 3% of the total bacteria in the rumen (Stevenson and Weimer, 2007; Mosoni *et al*., 2011). The primers used for quantification could be too specific to detect all members of these genera (Russell *et al*., 2009); however, recent evidence also suggest that microbes attached to plant substrates and the glycosyl hydrolases associated with them belong largely to bacteria that are phylogenetically distinct from these three species (Brulc *et al*., 2009; Hess *et al*., 2011).

Another area that is exclusively related to the activity of the microbiota, and for which more information is needed, is the production of enteric methane. Methane is produced by a specialised group of microorganisms, the methanogenic archaea (Liu and Whitman, 2008). The metabolic function of this group is known, the methanogenic community has a limited diversity compared with that of bacteria, and their numbers can be estimated at different rumen locations and followed in time (Janssen and Kirs, 2008). In addition, the genome sequence of the common rumen methanogen, *Methanobrevibacter ruminantium*, is available (Leahy *et al*., 2010) and sequencing of the genomes of other rumen methanogens is under way (see 'Rumen methanogen genome sequencing projects subsection'). However, in spite of all this information, methanogenesis in the rumen cannot be correlated to the number of methanogens (Yanez-Ruiz *et al*., 2008;

Mosoni *et al*., 2011; Popova *et al*., 2011) or unequivocally assigned to a particular community structure (Zhou *et al*., 2010; Morgavi *et al*., 2012). Methane production is intimately associated with the concentration of hydrogen and the interactions of methanogens with other microbes producing and consuming hydrogen in the rumen (Janssen, 2010; Morgavi *et al*., 2010).

The metabolism of complex carbohydrates and pathways leading to hydrogen disposal are two salient characteristics that have been identified in gut microbiomes (Brulc *et al*., 2009; Qin *et al*., 2010; Arumugam *et al*., 2011; Hess *et al*., 2011; Muegge *et al*., 2011). These two activities are linked as the use of plant structural carbohydrates and the removal of hydrogen are both significant elements in the efficient use of forage feed resources and the optimisation of fermentation in the anaerobic conditions of the rumen (Wolin *et al*., 1997). In addition, hydrogen is the central element influencing enteric methane production (Janssen, 2010). The incomplete knowledge of the microbiological controls of plant fibre degradation and methane production highlights the importance of an integrated approach to understand the function of the community and assess the importance of its members to the productivity of the rumen and the host animal. The application of second-generation sequencing technologies, the already-available single molecule sequencing technology (Schadt *et al*., 2010) and the future advent of fourth generation sequencing technologies with ever decreasing costs, longer reads and faster outputs (Perkel, 2011) makes possible the task of exploring the intricacy of the rumen using a systems biology approach (Raes and Bork, 2008). Raes and Bork stated that such an approach requires a comprehensive catalogue of the species and particularly the genes present in the system, along with an understanding of how these species/genes interact and fluctuate in time and space. In the following sections, we will summarise the main information available for the rumen ecosystem and propose areas where data have to be obtained to fill the gaps.

Rumen microbial diversity

The evolution of the study of rumen microbial diversity is similar to that of other microbial ecosystems, moving from culture-based and microscopic observations to the use of culture-independent, molecular techniques. The small-subunit ribosomal RNA gene (*rrs*) is the most commonly used target for characterising this diversity. For methanogens, the methyl coenzyme-M reductase (mcrA) gene of the methanogenesis pathway is also a phylogenetic marker (Luton *et al*., 2002) that has been used in rumen studies (Denman *et al*., 2007). Recently, the diversity of rumen methanogens was investigated using the gene encoding type II chaperonins (Chaban and Hill, 2012). Less universally distributed genes, indicative of certain functions, have also been used to study the parts of the microbial community. For example, the genes encoding the formyltetrahydrofolate synthetase (*fhs*) and the acetyl-CoA synthase (*acs*B) have been used to survey the range of potential homoacetogens in the rumen (Gagen *et al*., 2010; Henderson *et al*., 2010).

Monitoring particular microbial species using PCR-based amplification of *rrs* generates information on the presence and density of the target microbes in the rumen. Temporal and spatial changes in *rrs* copy numbers induced by different conditions can then be associated with biochemical and functional parameters of the ecosystem. This approach can provide useful information for evaluating the effect of dietary treatments on the rumen microbiome, for example, the changes associated with acidogenic diets in dairy cows (Khafipour *et al*., 2009). However, the species for which primers are available are generally those that can be cultured and these are not necessarily those that are most abundant. Indeed, it has been shown that most rumen bacterial species previously thought to be important represent only a small proportion of the total bacterial community (Stevenson and Weimer, 2007). In addition, there is no evidence that the ability to cultivate a given species is correlated with its functional importance in the ecosystem. The growth requirements of microbes and their ability to make the transition from their natural system to the laboratory are dependent on the intrinsic characteristics of a particular isolate. Recently, the application of simple existing methods has shown that a wider variety of rumen microbes can be cultured than previously realised (Kenters *et al*., 2010; Koike *et al*., 2010), opening up the possibility that the pure culture approach is still useful for understanding the properties of individual taxa. In particular, controlled experimentation to test hypotheses relating to gene function and microbial response to stimuli is greatly simplified when cultures are available. Rather than seeing molecular biological tools that allow community structure analyses and metagenomic investigations as invalidating culture-based investigations, we should use each approach for what it is best suited.

The use of fingerprinting techniques, of which the more commonly used in the rumen are single-strand conformation polymorphism (Michelland *et al*., 2009), denaturing gradient gel electrophoresis (DGGE; Kittelmann and Janssen, 2011; Popova *et al*., 2011), restriction fragment length polymorphism (RFLP) and its variant terminal-RFLP (t-RFLP; Khafipour *et al*., 2009; Yáñez-Ruiz *et al*., 2010) and automatic ribosomal intergenic spacer analysis (Sundset *et al*., 2009; Welkie *et al*., 2010), can provide more information on the structure of the rumen microbiome. These methods using a non-targeted approach are able to pick up differences and similarities in the community of different microbial groups in the rumen associated with host ruminant species, diet and feed efficiency (Larue *et al*., 2005; Sadet *et al*., 2007; Guan *et al*., 2008 and references above; Suen *et al*., 2011) but they do not provide direct sequence information. In addition, the number of bands or peaks detected by these techniques is several orders of magnitude lower than the estimated diversity in the rumen (Hess *et al*., 2011; Kim *et al*., 2011). As for all PCR-based techniques, the selection of primers and inherent problems during amplification can distort the diversity profiles obtained (Kanagawa, 2003;

Huws *et al*., 2007). In spite of these shortcomings, fingerprinting techniques will continue to be used as they provide a quick snapshot of the microbiota. They can show that there are differences between various treatments, but they cannot be used to catalogue the 'part list' necessary for a system biology approach (Raes and Bork, 2008).

Microbial diversity is better represented by the construction of libraries of conserved phylogenetic marker genes, commonly *rrs*, and sequencing them, which has allowed the accumulation of a vast number of rumen-derived sequences (Kim *et al*., 2011). The number of clones analysed per study increased exponentially with the refinement and decreasing cost of the Sanger DNA sequencing technique, for example, from 133 sequences examined by Whitford *et al*. (1998) in one of the first studies using clone libraries in the rumen to 11 171 produced by Durso *et al*. (2010) from cattle faecal samples. However, the production of data has been revolutionised by the availability of second-generation sequencing techniques, particularly pyrosequencing. This technique, although it has the limitations inherent to the PCR as stated above, can produce thousands of sequence tags in a single run at a fraction of the cost of traditional dye-terminator sequencing. The use of bar coding allows the use of multiple samples in the same run when applied to a single gene such as *rrs* (Tringe and Hugenholtz, 2008) and makes the technique even more affordable. The *rrs* has become the phylogenetic marker *par excellence*, with a vast coverage of sequences spanning all phyla known and from which inferences on the classification of novel sequences can be made (Head *et al*., 1998). In this paper, we will not discuss all the advantages and disadvantages of using *rrs* for phylogenetic classification of microbes, but there are a few aspects that should not be overlooked as they can affect diversity estimation. One of them is the presence of multiple rRNA operons in bacteria and archaea and the fact that not all operons present in a microbial genome are identical. This can lead to an overestimate of the total diversity, which has been calculated to be about 2.5-fold (Acinas *et al*., 2004). In addition, when quantification is performed, microbes that have high copy numbers of rRNA operons will be overestimated to the detriment of species possessing fewer copies than the community average (Crosby and Criddle, 2003). Other aspects, more on the technical side, are that primer selection may affect amplification of the community differently and the error attributed to the pyrosequencing technique itself that can lead to an overestimation of taxon abundance. If not corrected, this can be as much as 35% of the sequences (Gomez-Alvarez *et al*., 2009; Quince *et al*., 2009). Recently, the overestimation of taxa was illustrated when a single genome generated hundreds of different sequence types, leading to recommendations of strict quality filtering and careful application of sequence difference cut-offs for grouping sequences into operational taxa (Pukall *et al*., 2009; Purushe *et al*., 2010).

We have compared a pyrosequencing approach with study rumen fungi with parallel clone libraries of the internal transcribed spacer (ITS-1) gene region and found that, in general, the same community pattern is found using both techniques (S. Kittelmann *et al*., unpublished). In addition, comparison with DGGE also revealed that communities could be differentiated by any of the three methods (DGGE, clone libraries, pyrosequencing). Rarer members of the community were more easily detected in the larger pyrosequencing datasets. This study also revealed that more than 27% of reference rumen fungal sequences retrieved from GenBank was misnamed at the genus level, which confounds efforts to compare communities.

In-depth *rrs* sequencing shows that the community structure of rumen bacteria is affected by changes in diet composition (Callaway *et al*., 2010; Pitta *et al*., 2010). The community structure in the lower gastrointestinal tract is, as expected, different from that of the rumen (Callaway *et al*., 2010) but also greatly influenced by diet (Callaway *et al*., 2010; Durso *et al*., 2010; Shanks *et al*., 2011). The complexity of the feed seems to favour diversity. Rumen communities associated with bermudagrass diets, rich in structural carbohydrates and secondary compounds, were more diverse than those associated with growing wheat forage (Pitta *et al*., 2010), and the presence of highly degradable carbohydrates in the diet such as starch decreased bacterial diversity in faeces (Shanks *et al*., 2011). As in other mammalian gastrointestinal microbiomes (Ley *et al*., 2008b), the predominant phyla are the *Bacteroidetes* and the *Firmicutes* (Brulc *et al*., 2009; Callaway *et al*., 2010; Durso *et al*., 2010). The microbial diversity in the gastrointestinal tract of ruminants is estimated to be higher than that of humans (Eckburg *et al*., 2005; Brulc *et al*., 2009; Durso *et al*., 2010), with a few predominant genera (20 to 25), representing up to 90% of the total number of sequences (Callaway *et al*., 2010; Durso *et al*., 2010). Depending on the anatomical site, rumen or caecum, and probably also on the technical approach, the relative proportion of genera may change. However, most studies identified *Prevotella* as an important, if not the predominant, genus in the community. These studies also confirm that the cultivated cellulolytic bacterial genera *Ruminococcus* and *Fibrobacter* are not among the most abundant members of the community, but detected instead the presence of various other fibre-degrading genera (Brulc *et al*., 2009; Pitta *et al*., 2010).

Deep sequencing produces useful and extensive coverage of the microbial diversity, allowing us to identify the prevalent 'core' members of the community but also 'rare' community members that could be associated with feeding practices (Shanks *et al*., 2011). The rare members are otherwise not detectable using less sensitive techniques (Pedros-Alio, 2006). Notwithstanding, this expanded *rrs* dataset generally confirms information that was suggested by fingerprinting approaches or shallow sequencing of clone libraries. The *rrs* gene sequencing approach is a good method for making an initial phylogenetic classification of novel, not yet cultured microbes down to the genus level. However, the information contained in this single gene does not have enough resolution to define a species (Rosselló-Mora and Amann, 2001). The use of an expanded set of marker genes gives a better qualitative and quantitative picture of natural communities than *rrs*-based

approaches (von Mering *et al*., 2007). Importantly, these authors also determined that there is no automatic parallel between phylogenetic similarity and microbial phenotype or function. Microbes that are phylogenetically close can often have different functions and metabolic characteristics (Achenbach and Coates, 2000; von Mering *et al*., 2007), highlighting the importance of characterising the functional properties of the ecosystem.

Analysing communities at the phylogenetic level is useful in those cases where we can correlate function and identity, or where the aim is simply to show that there has been a change in community structure. However, because of the limitations in culturing rumen microbes, many marker genes indicate the presence of uncharacterised groups, and their physiologies have to be estimated because they are not yet known. By combining different approaches, investigation at a functional level (e.g. cellulose degradation, hydrogen metabolism) rather than a phylogenetic one is more readily possible. We can then start to (better) assign function, role and significance to differences in microbial community structure. Genome sequences from new cultures of rumen microbes offer the advantage that genes and phenotype can be experimentally linked. The range of genomes can be increased by selecting single cells. Using metagenomics, genome information from entire communities can be gathered and eventually reconstructed. Transcriptomics allows an estimation of expression of genes under the prevailing conditions, which in the end determines whether an activity or phenotype is actually present in the rumen, rather than just being detected as a gene or observed in laboratory cultures. Below, we discuss how these 'omics' approaches are starting to be applied to the rumen to improve our understanding of this important ecosystem.

Rumen microbial genome projects

Initial studies on individual genes of rumen microorganisms relied on retrieval of genes from libraries of genomic DNAs via functional screens or, more latterly, via PCR amplification of genes and their homologues. Gene expression was detected by heterologous expression in (mainly) *Escherichia coli* and sequencing was carried out on clones using Sanger dideoxy chain termination reactions separated in DNA sequencing gels and read from autoradiographs, 300 bp at a time (Sanger *et al*., 1977). Only the activities of those genes expressed in *E. coli* were detected, inevitably leading to a vast underestimate of the genes present. The first genes retrieved via these approaches encoded cellulases from *F. succinogenes* (Crosby *et al*., 1984). Subsequently, seven *F. succinogenes* genes encoding fibre-degrading enzymes were found using these conventional genetic approaches (Forsberg *et al*., 2000). This of course was a great advance, but when the genome of *F. succinogenes* S85 was sequenced, it revealed 104 open reading frames predicted to be involved in plant cell wall breakdown, including 33 cellulases, 24 xylanases and 14 carbohydrate esterases (Jun *et al*., 2007). The power of genome sequencing is therefore apparent from this case; several-fold more enzymes were identified in the *F. succinogenes* genome than that were known from all previous studies in this organism.

Advances in technology have seen genome sequencing become easier and more affordable. The pace of genome sequencing of individual rumen bacteria has continued to accelerate and has added significant impetus to rumen microbiology in recent years. There are currently 16 publicly available genome sequences from rumen microorganisms (Table 1), of which half are closed genomes. These sequencing projects have differing objectives, from discovering new genes encoding fibre-degrading activities for enhancing animal production or for feedstock depolymerisation for biofuel production, through to ruminant methane mitigation via interventions against methanogens. These projects are also adding functional 'omics' technologies to gain a better understanding of gene function in these organisms. Differential gene expression (transcriptomics) and protein expression (proteomics) or production of metabolites (metabolomics) are being used to confirm the functions of annotated genes and to test hypotheses around the role of 'unknown' or 'conserved hypothetical' genes. These initial rumen microbial genomes were produced from type strains that were maintained as single cultures for years. It has to be noted that these type strain might have some differences with strains evolving in their natural habitat, as their genetic makeup can change over generations (Papadopoulos *et al*., 1999) and mutations can arise induced by the culture media (Deng and Fong, 2011). In addition, variability within single species also exists (Hansen *et al*., 2011). Sequencing newly obtained isolates and single cells from the rumen will both determine the validity of type strains as models and inform the intra-species variation in the wild population.

The bacterial genome sequences completed to date have added a tremendous amount of gene sequences to the growing database of rumen microbial genes. The closed bacterial genomes alone have so far identified over 27 000 protein coding sequences, and this is likely to double in the near future with the closure of several draft genomes and the instigation of new rumen microbial genome projects. This knowledge is likely to contribute beneficially to aspects of ruminant digestive processes and may also positively influence ruminant products, for example, by suggesting how to increase conjugated linoleic acid levels in milk or meat. Structural carbohydrates form complex arrangements within the plant cell wall, and rumen microbes need a correspondingly elaborate array of enzymatic activities to degrade them. These enzymatic activities are found in a wide variety of phylogenetically distinct microbes, mainly in the bacteria, but also within fungi and to a lesser degree in protozoa. As indicated in Table 1, the majority of the current genome sequencing projects target cellulose- and hemicellulose-degrading bacteria, and one would expect the maximum benefit to ruminant production to occur by enhancing ruminal fibre degradation after enzymatic pretreatment of supplemental feeds. Other manipulations may become apparent from better knowledge of the microbes themselves.

Table 1 *Publically available genome sequences of rumen bacteria*

Organism	Contigs	Size (bp)	G + C%	Family	Sequencing centre[a]	Reference
Fibrolytic bacteria						
Butyrivibrio proteoclasticus B316[b]	1	3 555 059	40	*Lachnospiraceae*	AgResearch	Kelly *et al.* (2010)
	1	361 399	39			
	1	302 355	40			
	1	186 328	38			
Eubacterium cellulosolvens 6	107	3 260 436	48	*Lachnospiraceae*	JGI	
Prevotella bryantii $B_1$4	98	3 592 947	39	*Prevotellaceae*	NACGFRB	Purushe *et al.* (2010)
Prevotella ruminicola 23[b]	1	3 619 559	47	*Prevotellaceae*	NACGFRB	Purushe *et al.* (2010)
Fibrobacter succinogenes S85[b]	1	3 842 635	48	*Fibrobacteraceae*	JGI/NACGFRB	Suen *et al.* (2011)
Ruminococcus albus 7[b]	1	3 685 408	44	*Ruminococcaceae*	JGI	
	1	420 706	38			
	1	352 646	44			
	1	15 907	36			
	1	7420	42			
Ruminococcus albus 8	245	4 373 730	46	*Ruminococcaceae*	NACGFRB	
Ruminococcus flavefaciens FD-1	119	4 573 608	45	*Ruminococcaceae*	NACGFRB	Berg Miller *et al.* (2009)
Other bacteria						
Actinobacillus succinogenes 130Z[b]	1	2 319 663	44	*Pasteurellaceae*	JGI	
Desulfotomaculum ruminis DSM 2154	66	3 864 667	47	*Peptococcaceae*	JGI	
Desulfovibrio desulfuricans subsp. desulfuricans str. ATCC 27774[b]	1	2 873 437	58	*Desulfovibrionaceae*	JGI	
Basfia succiniciproducens MBEL55E[b]	1	2 314 078	42	*Pasteurellaceae*	KAIST	Hong *et al.* (2004)
Slackia heliotrinireducens DSM 20476[b]	1	3 165 038	60	*Coriobacteriaceae*	JGI	Pukall *et al.* (2009)
Treponema saccharophilum DSM 2985	56	3 453 897	53	*Spirochaetaceae*	JGI	
Wolinella succinogenes strain DSM 1740[b]	1	2 110 355	48	*Helicobacteriaceae*	MPI	Baar *et al.* (2003)
Megasphaera elsdenii DSM 20460	1	2 474 718	53	*Veillonellaceae*	BOKU	Marx *et al.* (2011)

[a]JGI, Joint Genome Institute; NACGFRB, North American Consortium for the Genomics of Fibrolytic Rumen Bacteria; KAIST, Korea Advanced Institute of Science and Technology; MPI, Max Planck Institute; BOKU,University of Natural Resources and Life Sciences, Vienna, Austria.
[b]Closed genomes.

What do the genomes of fibre-degrading rumen bacteria reveal about their strategies for polysaccharide degradation?
The different species of bacteria responsible for plant polysaccharide solubilisation each appear to have a different approach to fibre breakdown. *F. succinogenes* S85 appears to specialise as a cellulose degrader. Although its genome encodes a variety of enzymes capable of hydrolysing a wide range of plant cell wall polysaccharides, *F. succinogenes* lacks genes encoding transporter proteins or the enzymes involved in metabolising non-cellulose polysaccharides. Thus, it appears to encode enzymatic activities targeting non-cellulose polysaccharides only to clear its way to its primary substrate – cellulose. Only small amounts of extracellular cellulase activity can be detected in the supernatant from *F. succinogenes* cultures (Groleau and Forsberg, 1981), and adherence to solid cellulosic substrates appears to be a prerequisite for cellulose degradation (Kudo *et al.*, 1987; Weimer and Odt, 1995). *F. succinogenes* also lacks the cellulosome components such as dockerins, cohesions and scaffoldins that are found in *Clostridium thermocellum* and *R. flavefaciens* (Doi and Kosugi, 2004). Thus, it appears that cellulose degradation by *F. succinogenes* involves significant interaction between the cell and its enzymes (Lynd *et al.*, 2002) and it may well involve a cell-based, non-enzymatic process (Brumm *et al.*, 2011).

Ruminococcus species, on the other hand, encode a wide variety of enzymes predicted to be secreted from the cell and assembled into a highly organised enzyme scaffold, called a cellulosome, which mediates plant fibre degradation. The genes identified within the genome of *R. flavefaciens* FD-1 indicates that it can use a range of plant cell wall polysaccharides (Berg Miller *et al.*, 2009). It also confirms the presence of a multi-enzyme cellulosome complex in which enzymes are linked to a non-catalytic scaffold structure via dockerin domains. The cellulosome organisation in FD-1 is extremely complex and more than 200 dockerin-containing proteins have been identified from the draft genome sequence (Rincon *et al.*, 2010). Analysis of gene expression in this organism indicates that the type of substrate used by *R. flavefaciens* FD-1 can influence the enzymatic composition of the cellulosome.

The *R. albus* 7 genome is now closed and has been deposited on the US Department of Energy Joint Genome Institute website (http://genome.jgi-psf.org/rumal/rumal.info.html). Its genome is organised into a ~3.7 Mbp chromosome, two megaplasmids of 420 kbp and 352 kbp and two smaller plasmids of 15.9 kbp and 7.4 kbp. On the basis of our analysis of the *R. albus* 7 genome sequence, it contains far fewer dockerin-like proteins than *R. flavefaciens* FD-1. *R. albus* 7 has 29 dockerin-containing proteins, of which 25 are on the main

chromosome, three are on the 420 kbp plasmid and one is located on the 352 kbp plasmid. Several of these dockerin proteins contain polysaccharide-degrading domains. The genome sequence of another strain, *R. albus* 8, is less well annotated, but it seems to have ~40 proteins identified as containing Type I dockerins (http://www.jcvi.org/rumenomics). The smaller number of dockerin proteins presumably indicates that *R. albus* has a less well-developed cellulosome compared with *R. flavefaciens*. The draft *R. albus* 8 genome sequence has been used previously to identify proteins involved in cellulose adhesion and degradation (Devillard *et al.*, 2004). Compared with the wild-type strain, mutant strains of *R. albus* 8 that are defective in cellulose degradation lack two large-molecular-weight proteins, which were identified as endocellulases Cel48A and Cel9B. A Type III glutamine synthetase, an enzyme involved in the assimilation of nitrogen in *R. albus*, has also been shown to be important in cells growing under low-nitrogen conditions (Amaya *et al.*, 2005).

Butyrivibrio and *Pseudobutyrivibrio* species have come to prominence recently because of their strong xylan-degrading abilities and the prevalence of their *rrs* genes in pyrotag sequence libraries derived from rumen contents (Attwood *et al.*, 2004; C.D. Moon, personal communication). The recent genome sequencing of *Butyrivibrio proteoclasticus* B316 has identified a large number of genes involved in the use of plant polysaccharides, particularly hemicelluloses (Kelly *et al.*, 2010). These included glycosyl hydrolases (mainly from glycoside hydrolase (GH) families 2, 3, 5, 10, 13 and 43), carbohydrate esterases attacking ferulic acid and acetylxylan linkages and a variety of pectate lyases. Many of the *B. proteoclasticus* B316 GHs are annotated as belonging to various classes of xylanases, xylosidases or arabinosidases, consistent with its ability to grow well on xylan. However, despite *B. proteoclasticus* B316 being unable to grow on cellulose as a carbon source, two cellulase genes were identified. Proteomic analysis has shown that at least one of these cellulases is expressed constitutively during growth on xylan or xylose (Dunne *et al.*, 2007). Although the exact activities of these enzymes have not been defined, it seems likely that they contribute by improving access to hemicelluloses and pectins, which appear to be the preferred substrates of *B. proteoclasticus* B316. An interesting characteristic of the *B. proteoclasticus* enzymes involved in polysaccharide breakdown is that, based on signal peptide analysis, most are predicted to be found within the cell. The predicted enzyme localisation suggests that a limited number of secreted enzymes generate a variety of complex oligosaccharides that are then transported into the cell for further breakdown and metabolism. The clustering of genes encoding intracellular polysaccharide-degrading enzymes with genes encoding for transporters, transcriptional regulators and environmental sensors in polysaccharide utilisation loci (Kelly *et al.*, 2010) supports this model. Gene expression data also suggest co-expression of genes within these clusters when the *B. proteoclasticus* is grown on the insoluble substrate, xylan. Coordinated expression of polysaccharide-degrading enzymes with sugar transporters is likely to be a mechanism to reduce competition from saccharolytic microbes by limiting the amount of readily useable saccharides released into the surrounding environment. *B. proteoclasticus* has also been shown to adhere to plant material and this, combined with its extensive hemicellulose-degrading abilities, implies that it is capable of plant colonisation and breakdown in the rumen environment. Given that hemicellulose is predicted in most models of plant cell walls to surround cellulose, and thus one of the first plant polysaccharides encountered by invading bacteria, it is possible that *B. proteoclasticus* is an initial coloniser and degrader of plant material in the rumen, clearing the hemicellulose and allowing access to the cellulose by the primary cellulolytic microbes.

Rumen methanogen genome sequencing projects

Methane is an important greenhouse gas and emissions from agricultural sources account for ~40% of total anthropogenic methane, of which 25% is directly from enteric fermentation in livestock (Olivier *et al.*, 2005). There are several rumen methanogen genome sequencing projects either completed or underway that are improving our understanding of the diversity and metabolic capacity of methanogenic archaea in the rumen (Table 2).

M. ruminantium M1 was the first rumen methanogen to have its genome sequenced. This sequence information confirmed the hydrogenotrophic lifestyle of *M. ruminantium* M1, although gene expression studies indicated that formate may be an important substrate for methanogenesis during syntrophic growth with rumen bacteria (Leahy *et al.*, 2010). Short-chain alcohols were also predicted and were shown to stimulate growth on hydrogen, but not to support growth alone. A noticeable omission from the *M. ruminantium* genome was the lack of genes encoding the methyl coenzyme-M reductase II enzyme system (Mcr II), an isoenzyme of the Mcr I enzyme that is usually found in hydrogenotrophic methanogens. In other methanogens, the Mcr II system is differentially regulated during growth (Reeve *et al.*, 1997; Luo *et al.*, 2002) and is thought to mediate methane formation at high hydrogen partial pressures. In the rumen, methanogens depend on fermentative microbes to supply hydrogen, which is usually kept at very low concentrations (Janssen, 2010). The absence of a Mcr II system in *M. ruminantium* suggests that it is adapted for growth under low levels of hydrogen using the Mcr I system only.

As in the genomes of the human gut methanogens *Methanosphaera stadtmanae* (Fricke *et al.*, 2006) and *Methanobrevibacter smithii* (Samuel *et al.*, 2007), the *M. ruminantium* genome encodes large surface proteins that have features similar to bacterial adhesins. The role of such proteins is not known; however, in co-culturing experiments, genes encoding several of these adhesin-like proteins were upregulated and microscopic examination showed co-aggregation of *M. ruminantium* with *B. proteoclasticus* cells (Leahy *et al.*, 2010). The abundance of genes encoding such adhesin-like proteins in the *M. ruminantium* genome indicates a significant ability to modulate cell surface

Table 2 *Rumen methanogen genome sequencing projects*

Order	Genus	Species/(clade or closest relative)[a]	Strain[b]	Origin[c]	Status
Methanobacteriales	*Methanobrevibacter*	*ruminantium*	M1	Bovine	Closed
			YLM1	Ovine	Draft
			YE286	Bovine	Draft
		(*gottschalkii*)	SM9	Ovine	Closed
	Methanobacterium	*bryantii*	YE299	Bovine	Draft
		(*formicicum*)	BRM9	Bovine	Closed
	Methanobrevibacter	(*wolinii*)	ABM4	Ovine	Draft
	Methanosphaera	(*stadtmanae*)	3F5	Ovine	Draft
Methanosarcinales	*Methanosarcina*	(*barkeri*)	CM1	Bovine	Draft
TALC	Not assigned	(*Thermoplasma* sp.)	TALC	Bovine	Closed

TALC = Thermoplasmales-associated lineage C.
[a]Validly assigned species, or clade after Janssen and Kirs (2008) or closest cultured relative as determined by small subunit ribosomal RNA gene sequence similarity.
[b]Methanogen species sequenced as part of the following programmes: M1 and SM9, NZ Pastoral Greenhouse Gas Research Consortium; YLM1 and BRM9, NZ Ministry of Agriculture and Forestry Sustainable Land Management and Climate Change Fund; 3F5 and ABM4, NZ Agricultural Greenhouse Gas Research Centre; YE 286 and YE299, Commonwealth Scientific and Industrial Research Organisation-Queensland Department of Primary Industry and Beef Cooperative Research Centre; TALC, Commonwealth Scientific and Industrial Research Organisation, Livestock Industries.
[c]All strains were originally isolated from rumen contents, except ABM4, which was isolated from ovine abomasal contents.

topology and point to a likely role in mediating close associations with hydrogen producing bacteria, fungi and protozoa. In addition to adhesin-like proteins, *M. ruminantium* encodes more than 50 genes involved in the synthesis and export of exopolysaccharides. This is consistent with previous reports that *M. ruminantium* produces a capsule layer (Kandler and König, 1978; Kandler and König, 1985).

An unexpected finding from the *M. ruminantium* genome sequence was the discovery of a methanogen prophage sequence, designated φmru (Attwood *et al*., 2008; Leahy *et al*., 2010). Analysis of the φmru prophage sequence identified genes encoding cell lysis function. One of these genes, encoding an endoisopeptidase designated PeiR, was sub-cloned and expressed in *E. coli* and shown to mediate *M. ruminantium* cell lysis *in vitro* (Leahy *et al*., 2010). The discovery of two non-ribosomal peptide synthetases (NRPS) genes was also surprising, as they were the first to be reported in an archaeal genome. Non-ribosomal peptides (NRPs) have been attributed a wide range of activities, notably toxic, antimicrobial and iron chelating, but in this case their functions are unknown.

Recently, the fields of rumen genomics and metagenomics have begun to overlap with the reconstruction of individual bacterial genomes from metagenomic data. The deep metagenomic sequencing study by Hess *et al*. (2011) was able to initially assemble 179 092 scaffolds out of the 268 Gbp of metagenomic sequence. After further analyses to verify scaffold integrity, 26 042 scaffolds greater than 10 kbp were identified. These validated scaffolds were binned by tetranucleotide frequencies and read coverage, giving 446 genome bins. After assignment to the closest phylogenetic order, each genome bin was compared with minimal set of core genes derived from all available reference genomes of that order. This produced 15 genome bins, which contained between 60% and 93% of their respective core genes. Further gene copy number analysis was carried out on the genes that were present only in single copy in all reference genomes of the respective phylogenetic order. These analyses suggested that near-complete draft genomes were successfully assembled. Genome sequence data from individual uncultured microbial cells isolated directly from the same complex rumen community were obtained using single-cell isolation techniques. Following whole-genome amplification, their DNAs were screened for *rrs* sequences and one of the single cells analysed matched to genome bin APb, one of the largest bins assembled from metagenomic data. The *rrs* sequence indicated that the organism was related to *Butyrivibrio fibrisolvens*. We have subsequently analysed the bin APb genome (2.41 Mbp) and found that it is a close match to the draft genome of a *Butyrivibrio hungatei* strain (3.37 Mbp, N. Palevich *et al*., unpublished).

The quest for functions through metagenomics

Metagenomics is the culture-independent genomic analysis of microbial communities and comprises the functional and sequence-based analysis of the collective microbial genomes contained in an environmental sample. Its application started using the functional screening approach and was mainly used to discover novel biological activities in complex ecosystems (Handelsman *et al*., 1998). With the development of high-throughput DNA sequencing technologies, this functional metagenomic approach has been complemented by sequence-based metagenomics, where metabolic activities are identified by comparison with various databases. As for any molecular biology technique, care in the steps preceding the application of metagenomics protocols, for example sampling procedure, sample handling and nucleic acid extraction, is crucial for obtaining good and reliable data.

Table 3 *Plant polysaccharide-degrading enzymes identified from rumen metagenomic libraries*

CAZy family	Activity (number)	Animal	Reference
GH3	β-glucosidase	Cow	Wang *et al.* (2009)
GH3	β-xylosidase	Cow	Shedova *et al.* (2009)
GH3	Unspecified (15)	Cow	Hess *et al.* (2011)
GH5	Endo-glucanase (7)	Cow	Ferrer *et al.* (2005)
GH5	Cellodextrinase (2)	Buffalo	Duan *et al.* (2009)
GH5	Endo-glucanase (12)	Buffalo	Duan *et al.* (2009)
GH5	Endo-β-1,4-glucanase	Buffalo	Liu *et al.* (2009b)
GH5	Endo-β-1,4-glucanase	Cow	Shedova *et al.* (2009)
GH5	Endo-β-1,4-glucanase (2)	Cow	Wang *et al.* (2009)
GH5	Endo-xyloglucanase	Cow	Wong *et al.* (2010a)
GH5	Exo-xyloglucanase	Cow	Wong *et al.* (2010b)
GH5	Unspecified (27)	Cow	Hess *et al.* (2011)
GH5	Endo-glucanase/xylanase	Yak	Chang *et al.* (2011)
GH5/GH26	Glucanase/mannanase/xylanase	Cow	Palackal *et al.* (2007)
GH5/GH26	Unspecified (1)	Cow	Hess *et al.* (2011)
GH8	Unspecified (2)	Cow	Hess *et al.* (2011)
GH9	Unspecified (20)	Cow	Hess *et al.* (2011)
GH10	Unspecified (21)	Cow	Hess *et al.* (2011)
GH13	Cyclomaltodextrinase	Cow	Ferrer *et al.* (2007)
GH26	Endo-glucanase (2)	Cow	Ferrer *et al.* (2005)
GH26	Mannanase (2)	Buffalo	Duan *et al.* (2009)
GH26	Unspecified (1)	Cow	Hess *et al.* (2011)
GH43	Exo-α-1,5-L-arabinase	Cow	Wong *et al.* (2008)
GH43	Endo-α-1,5-L-arabinase	Cow	Wong *et al.* (2009)
GH43	Arabinosidase/xylosidase (3)	Cow	Zhao *et al.* (2010)
GH43	Endo-xylanase	Cow	Zhao *et al.* (2010)
GH48	Unspecified	Cow	Hess *et al.* (2011)
GH57	α-amylase	Cow	Zhao *et al.* (2010)
CE6	Acetyl-xylan esterase	Cow	Lopez-Cortes *et al.* (2007)
–	Unspecified esterase (11)	Cow	Ferrer *et al.* (2005)
–	Unspecified esterase (2)	Cow	Zhao *et al.* (2010)

Functional metagenomics

Because the rumen has evolved to be an efficient and complex lignocellulose degradation system, this process has been the focus of metagenomic studies aimed at determining and capturing the diversity of enzyme activities present. Functional metagenomics has been used to identify hydrolytic enzymes of biotechnological interest using specific substrates, particularly enzymes involved in the deconstruction of structural plant polysaccharides. This approach was pioneered by Ferrer *et al.* (2005), who detected 9 endoglucanases, 12 esterases and 1 cyclodextrinase from a dairy cow rumen metagenomic library. Subsequently, several groups have used metagenomic libraries to isolate specific polysaccharide-degrading enzymes from the rumen (Table 3 and Supplementary Table 1). This approach is dependent on the availability of suitable bioassays for the activities of interest, and to date most attention has been focussed on cellulose and hemicellulose degradation mediated by enzymes belonging to GH family 5. Rumen metagenome libraries have also been used to screen for other bioactivities, including novel lipases (Liu *et al.*, 2009a; Bayer *et al.*, 2010), polyphenol oxidase (Beloqui *et al.*, 2006) and an enzyme capable of degrading 3,5,6-trichloro-2-pyridinol, a degradation product of the organophosphorus insecticide chlorpyrifos (Math *et al.*, 2010). In addition, cysteine phytases have been isolated by screening ruminal genomic DNA from cows and goats using degenerate primer sets (Huang *et al.*, 2010).

Functional metagenomics has the potential to uncover new enzymes and metabolic pathways in the rumen if innovative strategies for screening are developed. The enzymatic machinery necessary to hydrolyse structural plant polysaccharides is an obvious target; however, although a new carbohydrate-binding domain has been reported (Duan *et al.*, 2009), the catalytic domains observed to date all fall into known families. Other potential targets include the biotransformation of compounds of interest for the animal or the consumer, such as antioxidants and conjugated linoleic acids or the detoxification of plant toxins, mycotoxins and xenobiotics. Other strategies used in functional metagenomic screening are the heterologous complementation of host strains and mutants and induction of reporter genes (Simon and Daniel, 2011). A drawback of this technique is that expression microbial hosts do not always express the proteins from other taxonomic groups in a functional form.

Sequence-based metagenomics

The exploration of mammalian gut microbiomes through sequence-based metagenomics is a recent development and reflects the increasing attention being given to the role that gut microbial communities have on health and disease of the host. The impetus comes from large sequencing projects on human gut microbial communities backed by the European Union (www.metahit.eu) and the National Institutes of Health in the United States of America (Human Microbiome Project, https://commonfund.nih.gov/hmp). We are aware of many ongoing projects using a metagenomic approach on the rumen microbiome, but there are still few published reports. Although different approaches have been used, large sequence-based studies have begun to catalogue the genes present and investigate rumen processes involved in fibre degradation (Brulc *et al*., 2009). The value of exploring the rumen environment in this way has been shown by the Joint Genome Institute study focussing on the microbial community attached to switchgrass incubated in the dairy cow rumen (Hess *et al*., 2011; Table 3), and by the metatranscriptomic analysis of the rumen of the muskox, which focussed on anaerobic fungi and protozoa (Qi *et al*., 2011). Table 4 compares the distribution of the main GH families involved in plant polysaccharide degradation in these metagenomic/metatranscriptomic studies, and in the four rumen fibrolytic bacteria whose genomes have been completely sequenced. It is notable that GH families involved in cellulose degradation are well represented in the two cellulolytic bacteria (*F. succinogenes* and *R. albus*) and in the muskoxen microbial eukaryotic metatranscriptome, but are less prevalent in the other studies. The metagenomic profiles appear more similar to those shown by hemicellulose- and pectin-degrading organisms (*Butyrivibrio* and *Prevotella*). Overall the main GH activities found (GH43, GH3, GH2, GH13) correlate with those detected in metagenomes from a range of other microbial environments (Li *et al*., 2009).

It would be helpful if the sequence-based metagenomic data could be placed in a taxonomic context. Analysis of metagenomic data would be facilitated by a reference set of rumen microbial genome sequences, analogous to the catalogue of reference genomes from the human microbiome (The Human Microbiome Jumpstart Reference Strains, 2010), being used to underpin analysis of human gut metagenomic datasets. The value of reference genomes for this type of analysis is seen in the rapid progress that is being made in this research area (Arumugam *et al*., 2011). New software can now handle RNA sequence data without the need for reference genomes (Grabherr *et al*., 2011), but has yet to be tested in complex ecosystems such as the rumen. Notwithstanding, the availability of more genomes and a good-quality metagenome reference gene catalogue can only enhance the promises of metatranscriptomics to uncover functional traits. Coupled with the availability of cultures, this allows functions to be linked to genes in laboratory experiments using reference organisms.

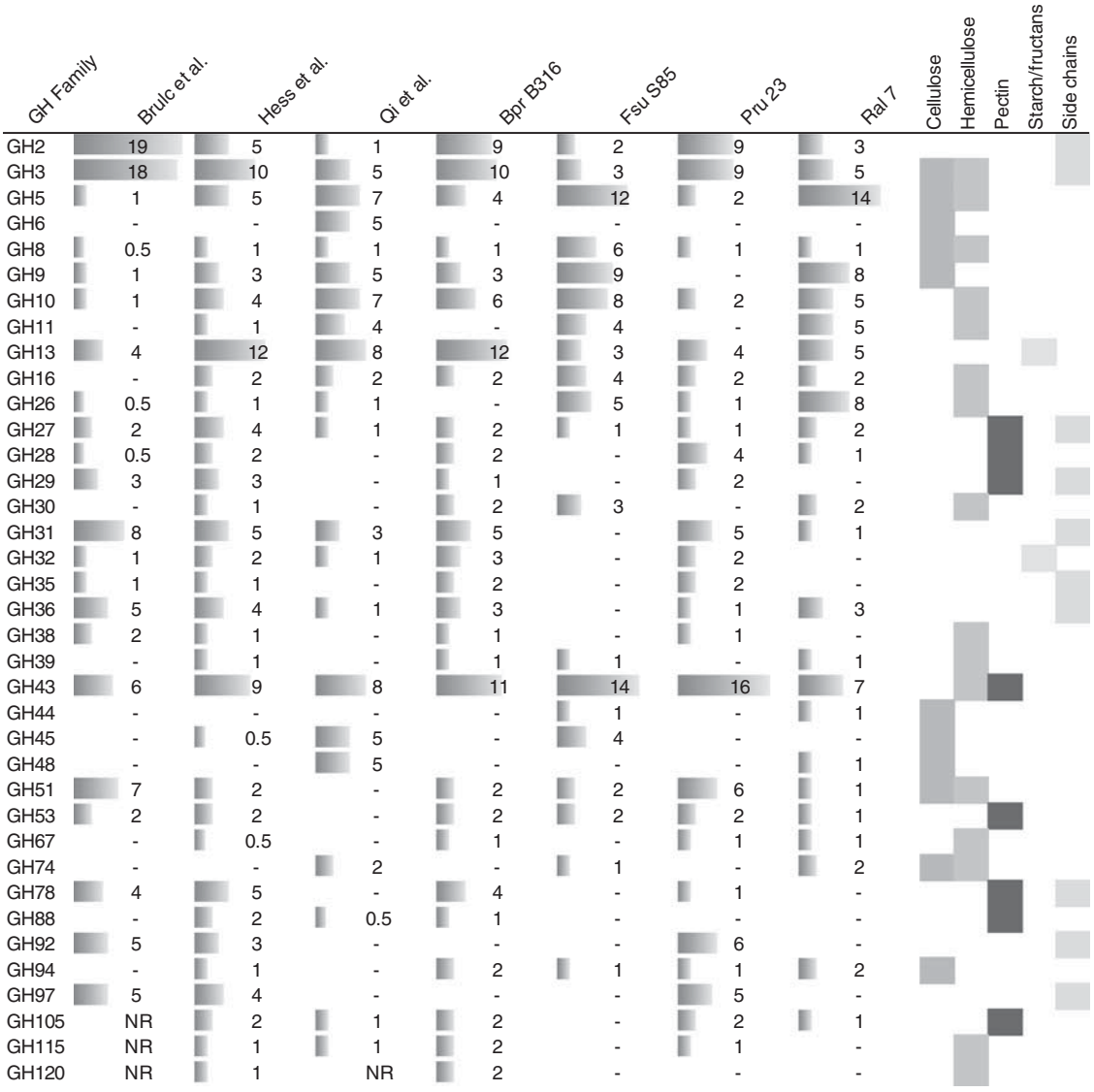
Linking genomics and metagenomics data to nutrition and other animal production features

Cataloguing the genes and functions of the microbiome is the step necessary for modelling and linking the rumen microbial metabolism to that of the host (Raes and Bork, 2008; Karlsson *et al*., 2011).

Traditionally, nutrition and production traits in ruminants were assessed without consideration of the rumen microbiome. At present, the inclusion of this 'component' of the superorganism is a great opportunity to improve our understanding of the digestion and biological characteristics of the host. As stated in previous sections, the rumen microbiome is complex, hugely diverse and its gene catalogue is incomplete. On the basis of ecological concepts, the taxonomical and functional diversity of the microbiome should provide high levels of redundancy and, consequently, stability (Konopka, 2009). Stability results in the maintenance of functions following changes in diet or other environmental stresses. However, there are many basic ecological questions not yet answered. Key community members and functions have to be conclusively defined in order to answer simple yet relevant questions, such as whether all numerically dominant rumen microbes are key members of the community? Perhaps key functions are carried out by rare members. Rare members constitute an 'insurance' to the host during changes in the environment, as has been described for some specialised microbes degrading plant toxins (Weimer, 1998), but many more functions may be hidden within members of the rare community (Pedros-Alio, 2006).

Recently, Muegge *et al*. (2011) showed that mammalian microbiomes have in common a large set of functions, implying that insights obtained from human and other animal studies can inform general concepts applied to ruminants. Microbial lineages remain associated with a given environment (von Mering *et al*., 2007) and those present in the gut are not usually found in other environments (Ley *et al*., 2008a). In addition, gut microbes are largely shared between mammals (Ley *et al*., 2008b), reinforcing the idea that some concepts are universally applicable across humans and livestock species. In humans, a minimal metagenome core for the proper function of the microbiome and the minimal set of genes required for a bacterium to inhabit the gut have been identified (Qin *et al*., 2010). In addition, within human hosts, a reduced number of microbial clusters, called enterotypes, were identified (Arumugam *et al*., 2011). These enterotypes are characterised by the prevalence of different trophic chains, reflecting distinctive capacities of the host to extract energy from foods, and this concept may be applicable to ruminants (Arumugam *et al*., 2011). Interestingly, not all important microbial functions were provided by dominant bacteria and Arumugam *et al*. (2011) suggest that functional biomarkers are more robust than phylogenetic biomarkers for identifying enterotypes associated with host phenotype characteristics. The complementary metabolic pathways provided by the microbiome influence the phenotype of animals and humans and make them react

Table 4 *Profile of the main GH families involved in plant polysaccharide degradation in the rumen. Values shown represent the percentage of the total GHs found in each species/metagenome/metatranscriptome*



GH Family	Brulc et al.	Hess et al.	Qi et al.	Bpr B316	Fsu S85	Pru 23	Ral 7	Cellulose	Hemicellulose	Pectin	Starch/fructans	Side chains
GH2	19	5	1	9	2	9	3					
GH3	18	10	5	10	3	9	5					
GH5	1	5	7	4	12	2	14					
GH6	-	-	5	-	-	-	-					
GH8	0.5	1	1	1	6	1	1					
GH9	1	3	5	3	9	-	8					
GH10	1	4	7	6	8	2	5					
GH11	-	1	4	-	4	-	5					
GH13	4	12	8	12	3	4	5					
GH16	-	2	2	2	4	2	2					
GH26	0.5	1	1	-	5	1	8					
GH27	2	4	1	2	1	1	2					
GH28	0.5	2	-	2	-	4	1					
GH29	3	3	-	1	-	2	-					
GH30	-	1	-	2	3	-	2					
GH31	8	5	3	5	-	5	1					
GH32	1	2	1	3	-	2	-					
GH35	1	1	-	2	-	2	-					
GH36	5	4	1	3	-	1	3					
GH38	2	1	-	1	-	1	-					
GH39	-	1	-	1	1	-	1					
GH43	6	9	8	11	14	16	7					
GH44	-	-	-	-	1	-	1					
GH45	-	0.5	5	-	4	-	-					
GH48	-	-	5	-	-	-	1					
GH51	7	2	-	2	2	6	1					
GH53	2	2	-	2	2	2	1					
GH67	-	0.5	-	1	-	1	1					
GH74	-	-	2	-	1	-	2					
GH78	4	5	-	4	-	1	-					
GH88	-	2	0.5	1	-	-	-					
GH92	5	3	-	-	-	6	-					
GH94	-	1	-	2	1	1	2					
GH97	5	4	-	-	-	5	-					
GH105	NR	2	1	2	-	2	1					
GH115	NR	1	1	2	-	1	-					
GH120	NR	1	NR	2	-	-	-					

GH = glycoside hydrolase; NR = not reported.
Data from the two bovine metagenome studies (Brulc *et al.*, 2009; Hess *et al.*, 2011), the muskoxen microbial eukaryotic metatranscriptomic study (Qi *et al.*, 2011), and the completed genome sequences of four representative fibrolytic rumen bacteria. These are *Butyrivibrio proteoclasticus* B316 (Kelly *et al.*, 2010), *Fibrobacter succinogenes* S85 (Suen *et al.*, 2011), *Prevotella ruminicola* 23 (Purushe *et al.*, 2010) and *Ruminococcus albus* 7. Information on their GH profiles was obtained from the CAZy database (http://www.cazy.org/). Analysis of GH families not included in the original publications was carried out as described by Warnecke *et al.* (2007).

differently to diet and drugs (Nicholson *et al.*, 2005; Li *et al.*, 2008; Claus *et al.*, 2011).

Understanding microbial processes

The interactions among genes and functions present in the catalogue have to be interpreted, and complementary approaches should be used to address fundamental questions of rumen microbial ecology. Some examples of topics where further information is needed are listed below. This information will be useful to modulate the capabilities of the microbiome for improving ruminant production.

The repercussions of the microbiome on the phenotype of the superorganism can be found at different levels, spanning from the single microbial population up to the microbial community and host–microbial interactions. The gastrointestinal environment is perpetually changing (diet, feed and water consumption, peristaltic action, etc.), and symbionts are constantly adapting to be at their optimal fitness (Dethlefsen *et al.*, 2007). Variants from the same species can arise through mutations or acquisition of genetic material through horizontal gene transfer (HGT). These variants will occupy different trophic niches and can coexist (Rosenzweig *et al.*, 1994). In humans producing methane, a single methanogen, *M. smithii*, is often found (Eckburg *et al.*, 2005). A single microbe for a unique function could be regarded as straightforward. However, there exists a great diversity

of strains differing in their use of formate and the presence of adhesin-like proteins, which hypothetically allow these variants to interact with different bacteria and occupy different niches (Hansen *et al*., 2011). Considering the complexity and selection pressure prevailing in the rumen, it is certain that multiple ecotypes of methanogens, fibrolytic microbes and other microbes exist and play a functional role in the microbiome.

The rumen microbiome is rich in microbial interactions between members, such as metabolic cooperation, synergism, predation, cell–cell signalling and structural organisation such as biofilms (McAllister *et al*., 1994; Hobson and Stewart, 1997; Erickson *et al*., 2002). Genomic information, bioinformatic tools for the phylogenetic assignment of sequences (Weber *et al*., 2011) and the application of techniques that can provide information on direct interactions between microbes in the environment, such as fluorescence *in situ* hybridization and single-cell sequencing, will provide invaluable insights of the ecosystem.

A fundamental agent of change and adaptation in microbial communities are phages (Clokie *et al*., 2011). Indeed, in gut metagenomes, viral sequences are important, and represent nearly 6% of the total sequences found in the human microbiome (Arumugam *et al*., 2011). In addition, they are vectors of HGT, for which there is ample evidence in gut bacterial genomes and metagenomes. The rumen is not an exception and bacteriophage and viral sequences have been described (Klieve and Bauchop, 1988; Dinsdale *et al*., 2008). The virome is a fundamental agent of change of the microbiome dynamics and should not be overlooked in the study of complex communities (Rohwer and Youle, 2011).

Future prospects

It is clear from the preceding discussion that a catalogue of rumen microbial genes and assignment of functions to these genes are required to link the rumen microbiome to nutritional and production practices in ruminants. What is needed is an initiative within the rumen microbial community, similar to the Human Microbiome Project (Turnbaugh *et al*., 2007), to produce a reference set of rumen microbial genomes that would support the analysis and comprehension of large metagenomic datasets. It would also help form testable hypotheses that could direct experimentation and lead to a better understanding of rumen biology. This could be achieved by sequencing the genomes of the available cultivated rumen bacteria and methanogenic archaea ($\sim$1000 cultures), together with representative cultures of rumen anaerobic fungi and protozoa. We have estimated, on the basis of an average bacterial and archaeal genome size of $\sim$3.5 and $\sim$2.5 Mbp, respectively, and sequencing coverage of $\sim$100$\times$, that 300 to 400 Gbp of Illumina sequencing would be required to achieve this goal. With each Illumina HiSeq paired end 100-bp sequence run now approaching 500 Gbp of sequence data above Q30, this is clearly attainable with current technology. The initial focus might be on culturable bacteria and archaea; however, new single-cell genome sequencing technologies may allow individual cells of uncultivated organisms to be sequenced if analysis of available metagenomic data consistently shows that taxa believed to be important in the rumen cannot be linked with cultured representatives. This would be a large project that would require significant funding, coordination and planning, but which would underpin and greatly enhance the understanding of the increasingly large amounts of metagenomics data that are being generated from the rumen.

Acknowledgements

The authors gratefully acknowledge the financial support for the methanogen genome sequencing and metagenomics projects from the following New Zealand funding agencies: the NZ Pastoral Greenhouse Gas Research Consortium, NZ Ministry for Agriculture and Forestry Sustainable Land Management Mitigation and Adaptation to Climate Change (SLMACC), New Economy Research Fund of the NZ Foundation for Research, Science and Technology and the NZ Agricultural Greenhouse Gas Research Centre. INRA's work in rumen metagenomics is funded by the French Ministry of Higher Education and Research and the French Ministry of Food, Agriculture and Fisheries.

Supplementary materials

For supplementary material referred to in this article, please visit http://dx.doi.org/doi:10.1017/S1751731112000419

References

Achenbach LA and Coates JD 2000. Disparity between bacterial phylogeny and physiology – comparing 16S rRNA sequences to assess relationships can be a powerful tool, but its limitations need to be considered. ASM News 66, 714–715.

Acinas SG, Marcelino LA, Klepac-Ceraj V and Polz MF 2004. Divergence and redundancy of 16S rRNA sequences in genomes with multiple rrn operons. Journal of Bacteriology 186, 2629–2635.

Amaya KR, Kocherginskaya SA, Mackie RI and Cann IKO 2005. Biochemical and mutational analysis of glutamine synthetase type III from the rumen anaerobe *Ruminococcus albus* 8. Journal of Bacteriology 187, 7481–7491.

Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto J-M, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Dore J, Weissenbach J, Ehrlich SD and Bork P 2011. Enterotypes of the human gut microbiome. Nature 473, 174–180.

Attwood GT, Cookson AC and Kelly WJ 2004. Genome sequencing of *Clostridium proteoclasticum*. Reproduction Nutrition Development 44 (Suppl 1), S20.

Attwood GT, Kelly WJ, Altermann EH and Leahy SC 2008. Analysis of the *Methanobrevibacter ruminantium* draft genome: understanding methanogen biology to inhibit their action in the rumen. Australian Journal of Experimental Agriculture 48, 83–88.

Baar C, Eppinger M, Raddatz G, Simon J, Lanz C, Klimmek O, Nandakumar R, Gross R, Rosinus A, Keller H, Jagtap P, Linke B, Meyer F, Lederer H and Schuster SC 2003. Complete genome sequence and analysis of *Wolinella succinogenes*. Proceedings of the National Academy of Sciences of the United States of America 100, 11690–11695.

Balter M 2010. In a cold snap, farmers turned to milk. Science 329, 1465.

Bayer S, Kunert A, Ballschmiter M and Greiner-Stoeffele T 2010. Indication for a new lipolytic enzyme family: isolation and characterization of two esterases from a metagenomic library. Journal of Molecular Microbiology and Biotechnology 18, 181–187.

Beloqui A, Pita M, Polaina J, Martinez-Arias A, Golyshina OV, Zumarraga M, Yakimov MM, Garcia-Arellano H, Alcalde M, Fernandez VM, Elborough K, Andreu JM, Ballesteros A, Plou FJ, Timmis KN, Ferrer M and Golyshin PN 2006. Novel polyphenol oxidase mined from a metagenome expression library of bovine rumen: biochemical properties, structural analysis, and phylogenetic relationships. Journal of Biological Chemistry 281, 22933–22942.

Benner SA, Caraco MD, Thomson JM and Gaucher EA 2002. Planetary biology–paleontological, geological, and molecular histories of life. Science 296, 864–868.

Bera-Maillet C, Devillard E, Cezette M, Jouany J-P and Forano E 2005. Xylanases and carboxymethylcellulases of the rumen protozoa *Polyplastron multivesiculatum*, *Eudiplodinium maggii* and *Entodinium* sp. FEMS Microbiology Letters 244, 149–156.

Berg Miller ME, Antonopoulos DA, Rincon MT, Band M, Bari A, Akraiko T, Hernandez A, Thimmapuram J, Henrissat B, Coutinho PM, Borovok I, Jindou S, Lamed R, Flint HJ, Bayer EA and White BA 2009. Diversity and strain specificity of plant cell wall degrading enzymes revealed by the draft genome of *Ruminococcus flavefaciens* FD-1. PLoS One 4, e6650.

Brulc JM, Antonopoulos DA, Miller ME, Wilson MK, Yannarell AC, Dinsdale EA, Edwards RE, Frank ED, Emerson JB, Wacklin P, Coutinho PM, Henrissat B, Nelson KE and White BA 2009. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. Proceedings of the National Academy of Sciences of the United States of America 106, 1948–1953.

Brumm P, Mead D, Boyum J, Drinkwater C, Gowda K, Stevenson D and Weimer P 2011. Functional annotation of *Fibrobacter succinogenes* S85 carbohydrate active enzymes. Applied Biochemistry and Biotechnology 163, 649–657.

Callaway TR, Dowd SE, Edrington TS, Anderson RC, Krueger N, Bauer N, Kononoff PJ and Nisbet DJ 2010. Evaluation of bacterial diversity in the rumen and feces of cattle fed different levels of dried distillers grains plus solubles using bacterial tag-encoded FLX amplicon pyrosequencing. Journal of Animal Science 88, 3977–3983.

Chaban B and Hill JE 2012. A 'universal' type II chaperonin PCR detection system for the investigation of Archaea in complex microbial communities. ISME Journal 6, 430–439.

Chang L, Ding M, Bao L, Chen Y, Zhou J and Lu H 2011. Characterization of a bifunctional xylanase/endoglucanase from yak rumen microorganisms. Applied Microbiology and Biotechnology 90, 1933–1942.

Chaucheyras-Durand F, Walker ND and Bach A 2008. Effects of active dry yeasts on the rumen microbial ecosystem: past, present and future. Animal Feed Science and Technology 145, 5–26.

Claus SP, Ellero SL, Berger B, Krause L, Bruttin A, Molina Jrm, Paris A, Want EJ, de Waziers I, Cloarec O, Richards SE, Wang Y, Dumas M-E, Ross A, Rezzi S, Kochhar S, Van Bladeren P, Lindon JC, Holmes E and Nicholson JK 2011. Colonization-induced host–gut microbial metabolic interaction. mBio 2, doi:10.1128/mBio.00271-10.

Clauss M, Hume ID and Hummel J 2010. Evolutionary adaptations of ruminants and their potential relevance for modern production systems. Animal 4, 979–992.

Clauss M, Müller K, Fickel J, Streich WJ, Hatt JM and Südekum KH 2011. Macroecology of the host determines microecology of endobionts: protozoal faunas vary with wild ruminant feeding type and body mass. Journal of Zoology 283, 169–185.

Clokie MRJ, Millard AD, Letarov AV and Heaphy S 2011. Phages in nature. Bacteriophage 1, 31–45.

Crosby LD and Criddle CS 2003. Understanding bias in microbial community analysis techniques due to rrn operon copy number heterogeneity. BioTechniques 34, 790–794, 796, 798 passim.

Crosby WL, Collier B, Thomas DY, Teather RM and Erfle JD 1984. Cloning and expression in *Escherichia coli* of cellulase genes from *Bacteroides succinogenes*. Proceedings of the 5th Canadian Bioenergy Research and Development Seminar, Barking, England, pp. 573–576.

Deng Y and Fong SS 2011. Laboratory evolution and multi-platform genome re-sequencing of the cellulolytic actinobacterium *Thermobifida fusca*. Journal of Biological Chemistry 286, 39958–39966.

Denman SE, Tomkins NW and McSweeney CS 2007. Quantitation and diversity analysis of ruminal methanogenic populations in response to the antimethanogenic compound bromochloromethane. FEMS Microbiology Ecology 62, 313–322.

Dethlefsen L, McFall-Ngai M and Relman DA 2007. An ecological and evolutionary perspective on human-microbe mutualism and disease. Nature 449, 811–818.

Devillard E, Goodheart DB, Karnati SKR, Bayer EA, Lamed R, Miron J, Nelson KE and Morrison M 2004. *Ruminococcus albus* 8 mutants defective in cellulose degradation are deficient in two processive endocellulases, cel48a and cel9b, both of which possess a novel modular architecture. Journal of Bacteriology 186, 136–145.

Diamond J 1997. Guns, germs, and steel: the fates of human societies. Norton, New York.

Diamond J 2002. Evolution, consequences and future of plant and animal domestication. Nature 418, 700–707.

Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, Furlan M, Desnues C, Haynes M, Li L, McDaniel L, Moran MA, Nelson KE, Nilsson C, Olson R, Paul J, Brito BR, Ruan Y, Swan BK, Stevens R, Valentine DL, Thurber RV, Wegley L, White BA and Rohwer F 2008. Functional metagenomic profiling of nine biomes. Nature 452, 629–632.

Doi RH and Kosugi A 2004. Cellulosomes: plant-cell-wall-degrading enzyme complexes. Nature Reviews Microbiology 2, 541–551.

Duan CJ, Xian L, Zhao GC, Feng Y, Pang H, Bai XL, Tang JL, Ma QS and Feng JX 2009. Isolation and partial characterization of novel genes encoding acidic cellulases from metagenomes of buffalo rumens. Journal of Applied Microbiology 107, 245–256.

Dunne JC, Li D, Kelly WJ, Leahy SC, Bond JJ, Jordan TW and Attwood GT 2007. Proteins involved in plant polysaccharide depolymerisation are differentially expressed in the *Clostridium proteoclasticum* secretome. In Proceedings of the Human Proteome Organisation 6th Annual World Congress 'Proteomics: From Technology Development of Biomarker Applications', 6–10 October 2007, Molecular and Cellular Proteomics, Seoul, Korea.

Durso LM, Harhay GP, Smith TP, Bono JL, Desantis TZ, Harhay DM, Andersen GL, Keen JE, Laegreid WW and Clawson ML 2010. Animal-to-animal variation in fecal microbial diversity among beef cattle. Applied and Environmental Microbiology 76, 4858–4862.

Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE and Relman DA 2005. Diversity of the human intestinal microbial flora. Science 308, 1635–1638.

Edwards JE, McEwan NR, Travis AJ and Wallace RJ 2004. 16S rDNA library-based analysis of ruminal bacterial diversity. Antonie van Leeuwenhoek 86, 263–281.

Egert M, de Graaf AA, Smidt H, de Vos WM and Venema K 2006. Beyond diversity: functional microbiomics of the human colon. Trends in Microbiology 14, 86–91.

Erickson DL, Nsereko VL, Morgavi DP, Selinger LB, Rode LM and Beauchemin KA 2002. Evidence of quorum sensing in the rumen ecosystem: detection of N-acyl homoserine lactone autoinducers in ruminal contents. Canadian Journal of Microbiology 48, 374–378.

FAO 2006. World agriculture: towards 2030/2050. Interim Report. Food and Agriculture Organization of the United Nations, Global Perspective Studies Unit, Rome.

FAO 2009. United Nations Food and Agriculture Organization. Retrieved June 1, 2011, from http://faostat.fao.org

Ferrer M, Beloqui A, Golyshina OV, Plou FJ, Neef A, Chernikova TN, Fernandez-Arrojo L, Ghazi I, Ballesteros A, Elborough K, Timmis KN and Golyshin PN 2007. Biochemical and structural features of a novel cyclodextrinase from cow rumen metagenome. Biotechnology Journal 2, 207–213.

Ferrer M, Golyshina OV, Chernikova TN, Khachane AN, Reyes-Duarte D, Santos VAPMD, Strompl C, Elborough K, Jarvis G, Neef A, Yakimov MM, Timmis KN and Golyshin PN 2005. Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. Environmental Microbiology 7, 1996–2010.

Flint HJ, Bayer EA, Rincon MT, Lamed R and White BA 2008. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. Nature Reviews Microbiology 6, 121–131.

Forsberg CW, Forano E and Chesson A 2000. Microbial adherence to plant cell wall and enzymatic hydrolysis. In Ruminant physiology digestion metabolism, growth and reproduction (ed. PB Cronje), pp. 79–98. CABI Publishing, Wallingford, UK.

Fricke WF, Seedorf H, Henne A, Kruer M, Liesegang H, Hedderich R, Gottschalk G and Thauer RK 2006. The genome sequence of *Methanosphaera stadtmanae* reveals why this human intestinal archaeon is restricted to methanol and H2 for methane formation and ATP synthesis. Journal of Bacteriology 188, 642–658.

Gagen EJ, Denman SE, Padmanabha J, Zadbuke S, Al Jassim R, Morrison M and McSweeney CS 2010. Functional gene analysis suggests different acetogen populations in the bovine rumen and tammar wallaby forestomach. Applied and Environmental Microbiology 76, 7785–7795.

Ghorbani GR, Morgavi DP, Beauchemin KA and Leedle JA 2002. Effects of bacterial direct-fed microbials on ruminal fermentation, blood variables, and the microbial populations of feedlot cattle. Journal of Animal Science 80, 1977–1985.

Gill HS, Shu Q and Leng RA 2000. Immunization with *Streptococcus bovis* protects against lactic acidosis in sheep. Vaccine 18, 2541–2548.

Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM and Nelson KE 2006. Metagenomic analysis of the human distal gut microbiome. Science 312, 1355–1359.

Gomez-Alvarez V, Teal TK and Schmidt TM 2009. Systematic artifacts in metagenomes from complex microbial communities. ISME Journal 3, 1314–1317.

Goodacre R 2007. Metabolomics of a superorganism. Journal of Nutrition 137, 259S–266S.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N and Regev A 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29, 644–652.

Griffith GW, Baker S, Fliegerova K, Liggenstoffer A, van der Giezen M, Voigt K and Beakes G 2010. Anaerobic fungi: Neocallimastigomycota. IMA Fungus: the Global Mycological Journal 1, 181–185.

Groleau D and Forsberg CW 1981. Cellulolytic activity of the rumen bacterium *Bacteroides succinogenes*. Canadian Journal of Microbiology 27, 517–530.

Gruby D and Delafond HMO 1843. Recherches sur des animalcules se développant en grand nombre dans l'estomac et dans les intestins, pedant la digestion des animaux herbivores et carnivores. Comptes Rendus de l'Académie des Sciences 17, 1304–1308.

Guan LL, Nkrumah JD, Basarab JA and Moore SS 2008. Linkage of microbial ecology to phenotype: correlation of rumen microbial ecology to cattle's feed efficiency. FEMS Microbiology Letters 288, 85–91.

Handelsman J, Rondon MR, Brady SF, Clardy J and Goodman RM 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chemistry & Biology 5, R245–R249.

Hansen EE, Lozupone CA, Rey FE, Wu M, Guruge JL, Narra A, Goodfellow J, Zaneveld JR, McDonald DT, Goodrich JA, Heath AC, Knight R and Gordon JI 2011. Pan-genome of the dominant human gut-associated archaeon, *Methanobrevibacter smithii*, studied in twins. Proceedings of the National Academy of Sciences of the United States of America 108, 4599–4606.

Head IM, Saunders JR and Pickup RW 1998. Microbial evolution, diversity, and ecology: a decade of ribosomal RNA analysis of uncultivated microorganisms. Microbial Ecology 35, 1–21.

Henderson G, Naylor GE, Leahy SC and Janssen PH 2010. Presence of novel, potentially homoacetogenic bacteria in the rumen as determined by analysis of formyltetrahydrofolate synthetase sequences from ruminants. Applied and Environmental Microbiology 76, 2058–2066.

Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T, Mackie RI, Pennacchio LA, Tringe SG, Visel A, Woyke T, Wang Z and Rubin EM 2011. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science 331, 463–467.

Hobson PN and Stewart CS 1997. The rumen microbial ecosystem. Chapman & Hall, London.

Hong SH, Kim JS, Lee SY, In YH, Choi SS, Rih J-K, Kim CH, Jeong H, Hur CG and Kim JJ 2004. The genome sequence of the capnophilic rumen bacterium *Mannheimia succiniciproducens*. Nature Biotechnology 22, 1275–1281.

Hooper LV 2004. Bacterial contributions to mammalian gut development. Trends in Microbiology 12, 129–134.

Huang H, Zhang R, Fu D, Luo J, Li Z, Luo H, Shi P, Yang P, Diao Q and Yao B 2010. Diversity, abundance and characterization of ruminal cysteine phytases suggest their important role in phytate degradation. Environmental Microbiology 13, 747–757.

Hume ID and Warner ACI 1980. Evolution of microbial digestion in mammals. In Digestive physiology and metabolism in ruminants (ed. Y Ruckebusch and P Thivend), pp. 665–684. MTP Press, Lancaster, UK.

Hungate RE 1966. The rumen and its microbes. Academic Press, New York.

Huws SA, Edwards JE, Kim EJ and Scollan ND 2007. Specificity and sensitivity of eubacterial primers utilized for molecular profiling of bacteria within complex microbial ecosystems. Journal of Microbiological Methods 70, 565–569.

Janssen PH 2010. Influence of hydrogen on rumen methane formation and fermentation balances through microbial growth kinetics and fermentation thermodynamics. Animal Feed Science and Technology 160, 1–22.

Janssen PH and Kirs M 2008. Structure of the archaeal community of the rumen. Applied and Environmental Microbiology 74, 3619–3625.

Jouany JP and Morgavi DP 2007. Use of 'natural' products as alternatives to antibiotic feed additives in ruminant production. Animal 1, 1443–1466.

Jun HS, Qi M, Ha JK and Forsberg CW 2007. *Fibrobacter succinogenes*, a dominant fibrolytic ruminal bacterium: transition to the post genomic era. Asian-Australasian Journal of Animal Sciences 20, 802–810.

Kanagawa T 2003. Bias and artifacts in multitemplate polymerase chain reactions (PCR). Journal of Bioscience and Bioengineering 96, 317–323.

Kandler O and König H 1978. Chemical composition of the peptidoglycan-free cell walls of methanogenic bacteria. Archives of Microbiology 118, 141–152.

Kandler O and König H 1985. Cell envelopes of archaebacteria. In The bacteria (ed. CR Woese and RS Wolfe), pp. 413–457. Academic Press, Inc., New York, NY.

Karlsson FH, Nookaew I, Petranovic D and Nielsen J 2011. Prospects for systems biology and modeling of the gut microbiome. Trends in Biotechnology 29, 251–258.

Kelly WJ, Leahy SC, Altermann E, Yeoman CJ, Dunne JC, Kong Z, Pacheco DM, Li D, Noel SJ, Moon CD, Cookson AL and Attwood GT 2010. The glycobiome of the rumen bacterium *Butyrivibrio Proteoclasticus* B316^T highlights adaptation to a polysaccharide-rich environment. PLoS One 5, e11942.

Kenters N, Henderson G, Jeyanathan J, Kittelmann S and Janssen PH 2010. Isolation of previously uncultured rumen bacteria by dilution to extinction using a new liquid culture medium. Journal of Microbiological Methods 84, 52–60.

Khafipour E, Li S, Plaizier JC and Krause DO 2009. Rumen microbiome composition determined using two nutritional models of subacute ruminal acidosis. Applied and Environmental Microbiology 75, 7115–7124.

Kim M, Morrison M and Yu Z 2011. Status of the phylogenetic diversity census of ruminal microbiomes. FEMS Microbiology Ecology 76, 49–63.

Kittelmann S and Janssen PH 2011. Characterization of rumen ciliate community composition in domestic sheep, deer, and cattle, feeding on varying diets, by means of PCR-DGGE and clone libraries. FEMS Microbiology Ecology 75, 468–481.

Klieve AV and Bauchop T 1988. Morphological diversity of ruminal bacteriophages from sheep and cattle. Applied and Environmental Microbiology 54, 1637–1641.

Koike S, Handa Y, Goto H, Sakai K, Miyagawa E, Matsui H, Ito S and Kobayashi Y 2010. Molecular monitoring and isolation of previously uncultured bacterial strains from the sheep rumen. Applied and Environmental Microbiology 76, 1887–1894.

Konopka A 2006. Microbial ecology: searching for principles. Microbe 1, 175–179.

Konopka A 2009. What is microbial community ecology? ISME Journal 3, 1223–1230.

Kudo H, Cheng K-J and Costerton JW 1987. Electron microscopic study of the methylcellulose-mediated detachment of cellulolytic rumen bacteria from cellulose fibers. Canadian Journal of Microbiology 32, 244–248.

Larue R, Yu Z, Parisi VA, Egan AR and Morrison M 2005. Novel microbial diversity adherent to plant biomass in the herbivore gastrointestinal tract, as revealed by ribosomal intergenic spacer analysis and rrs gene sequencing. Environmental Microbiology 7, 530–543.

Leahy SC, Kelly WJ, Altermann E, Ronimus RS, Yeoman CJ, Pacheco DM, Li D, Kong Z, McTavish S, Sang C, Lambie SC, Janssen PH, Dey D and Attwood GT 2010. The genome sequence of the rumen methanogen *Methanobrevibacter ruminantium* reveals new possibilities for controlling ruminant methane emissions. PLoS One 5, e8926.

Lederberg J 2000. Infectious History. Science 288, 287–293.

Ley RE, Lozupone CA, Hamady M, Knight R and Gordon JI 2008a. Worlds within worlds: evolution of the vertebrate gut microbiota. Nature Reviews Microbiology 6, 776–788.

Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R and Gordon JI 2008b. Evolution of mammals and their gut microbes. Science 320, 1647–1651.

Li LL, McCorkle SR, Monchy S, Taghavi S and van der Lelie D 2009. Bioprospecting metagenomes: glycosyl hydrolases for converting biomass. Biotechnology for Biofuels 2, 10.

Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H, Zhang Y, Shen J, Pang X, Zhang M, Wei H, Chen Y, Lu H, Zuo J, Su M, Qiu Y, Jia W, Xiao C, Smith LM, Yang S, Holmes E, Tang H, Zhao G, Nicholson JK, Li L and Zhao L 2008. Symbiotic gut

microbes modulate human metabolic phenotypes. Proceedings of the National Academy of Sciences of the United States of America 105, 2117–2122.

Liu K, Wang J, Bu D, Zhao S, McSweeney C, Yu P and Li D 2009a. Isolation and biochemical characterization of two lipases from a metagenomic library of China Holstein cow rumen. Biochemical and Biophysical Research Communications 385, 605–611.

Liu L, Feng Y, Duan C-J, Pang H, Tang J-L and Feng J-X 2009b. Isolation of a gene encoding endoglucanase activity from uncultured microorganisms in buffalo rumen. World Journal of Microbiology and Biotechnology 25, 1035–1042.

Liu Y and Whitman WB 2008. Metabolic, phylogenetic, and ecological diversity of the methanogenic Archaea. Annals of the New York Academy of Sciences 1125, 171–189.

Lopez-Cortes N, Reyes-Duarte D, Beloqui A, Polaina J, Ghazi I, Golyshina OV, Ballesteros A, Golyshin PN and Ferrer M 2007. Catalytic role of conserved HQGE motif in the CE6 carbohydrate esterase family. FEBS Letters 581, 4657–4662.

Luo H-W, Zhang H, Suzuki T, Hattori S and Kamagata Y 2002. Differential expression of methanogenesis genes of *Methanothermobacter thermoautotrophicus* (formerly *Methanobacterium thermoautotrophicum*) in pure culture and in cocultures with fatty acid-oxidizing syntrophs. Applied and Environmental Microbiology 68, 1173–1179.

Luton PE, Wayne JM, Sharp RJ and Riley PW 2002. The mcrA gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. Microbiology 148, 3521–3530.

Lynd LR, Weimer PJ, van Zyl WH and Pretorius IS 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiology and Molecular Biology Reviews 66, 506–577.

Mackie RI, McSweeney CS and Klieve AV 2002. Microbial ecology of the ovine rumen. In Sheep nutrition (ed. M Freer and H Dove), pp. 71–94. CABI Publishing, Wallingford, UK.

Martin C, Morgavi DP and Doreau M 2010. Methane mitigation in ruminants: from microbe to the farm scale. Animal 4, 351–365.

Marx H, Graf AB, Tatto NE, Thallinger GG, Mattanovich D and Sauer M 2011. Genome sequence of the ruminal bacterium *Megasphaera elsdenii*. Journal of Bacteriology 193, 5578–5579.

Math RK, Islam SMA, Cho KM, Hong SJ, Kim JM, Yun MG, Cho JJ, Heo JY, Lee YH, Kim H and Yun HD 2010. Isolation of a novel gene encoding a 3,5,6-trichloro-2-pyridinol degrading enzyme from a cow rumen metagenomic library. Biodegradation 21, 565–573.

McAllister TA, Bae HD, Jones GA and Cheng K-J 1994. Microbial attachment and feed digestion in the rumen. Journal of Animal Sciences 72, 3004–3018.

Michelland RJ, Monteils V, Zened A, Combes S, Cauquil L, Gidenne T, Hamelin J and Fortun-Lamothe L 2009. Spatial and temporal variations of the bacterial community in the bovine digestive tract. Journal of Applied Microbiology 107, 1642–1650.

Morgavi DP, Forano E, Martin C and Newbold CJ 2010. Microbial ecosystem and methanogenesis in ruminants. Animal 4, 1024–1036.

Morgavi DP, Martin C, Jouany J-P and Ranilla MJ 2012. Rumen protozoa and methanogenesis: not a simple cause–effect relationship. British Journal of Nutrition 107, 388–397.

Morrison M, Pope PB, Denman SE and McSweeney CS 2009. Plant biomass degradation by gut microbiomes: more of the same or something new? Current Opinion in Biotechnology 20, 358–363.

Mosoni P, Martin C, Forano E and Morgavi DP 2011. Long-term defaunation increases the abundance of cellulolytic ruminococci and methanogens but does not affect the bacterial and methanogen diversity in the rumen of sheep. Journal of Animal Science 89, 783–791.

Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A, Fontana L, Henrissat B, Knight R and Gordon JI 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332, 970–974.

Nagaraja TG, Newbold CJ, Van Nevel CJ and Demeyer DI 1997. Manipulation of ruminal fermentation. In Rumen microbial ecosystem (ed. PN Hobson and CS Stewart), pp. 523–632. Blackie Academic & Professional, London.

Nicholson JK, Holmes E and Wilson ID 2005. Gut microorganisms, mammalian metabolism and personalized health care. Nature Reviews Microbiology 3, 431–438.

Olivier J, Van Aardenne J, Dentener F, Pagliari V, Ganzeveld L and Peters J 2005. Recent trends in global greenhouse gas emissions: regional trends 1970–2000 and spatial distribution of key sources in 2000. Environmental Sciences 2, 81–99.

Palackal N, Lyon C, Zaidi S, Luginbühl P, Dupree P, Goubet F, Macomber J, Short J, Hazlewood G, Robertson D and Steer B 2007. A multifunctional hybrid glycosyl hydrolase discovered in an uncultured microbial consortium from ruminant gut. Applied Microbiology and Biotechnology 74, 113–124.

Papadopoulos D, Schneider D, Meier-Eiss J, Arber W, Lenski RE and Blot M 1999. Genomic evolution during a 10,000-generation experiment with bacteria. Proceedings of the National Academy of Sciences 96, 3807–3812.

Pedros-Alio C 2006. Marine microbial diversity: can it be determined? Trends in Microbiology 14, 257–263.

Perkel J 2011. Making contact with sequencing's fourth generation. BioTechniques 50, 93–95.

Pitta D, Pinchak W, Dowd S, Osterstock J, Gontcharova V, Youn E, Dorton K, Yoon I, Min B, Fulford J, Wickersham T and Malinowski D 2010. Rumen bacterial diversity dynamics associated with changing from bermudagrass hay to grazed winter wheat diets. Microbial Ecology 59, 511–522.

Popova M, Martin C, Eugène M, Mialon MM, Doreau M and Morgavi DP 2011. Effect of fibre- and starch-rich finishing diets on methanogenic Archaea diversity and activity in the rumen of feedlot bulls. Animal Feed Science and Technology 166–167, 113–121.

Pukall R, Lapidus A, Nolan M, Copeland A, Del Rio TG, Lucas S, Chen F, Tice H, Cheng JF, Chertkov O, Bruce D, Goodwin L, Kuske C, Brettin T, Detter JC, Han C, Pitluck S, Pati A, Mavrommatis K, Ivanova N, Ovchinnikova G, Chen A, Palaniappan K, Schneider S, Rohde M, Chain P, D'Haeseleer P, Goker M, Bristow J, Eisen JA, Markowitz V, Kyrpides NC, Klenk HP and Hugenholtz P 2009. Complete genome sequence of *Slackia heliotrinireducens* type strain (RHS 1(T)). Standards in Genomic Sciences 1, 234–241.

Purushe J, Fouts DE, Morrison M, White BA, Mackie RI, Coutinho PM, Henrissat B and Nelson KE 2010. Comparative genome analysis of *Prevotella ruminicola* and *Prevotella bryantii*: insights into their environmental niche. Microbial Ecology 60, 721–729.

Qi M, Wang P, O'Toole N, Barboza PS, Ungerfeld E, Leigh MB, Selinger LB, Butler G, Tsang A, McAllister TA and Forster RJ 2011. Snapshot of the eukaryotic gene expression in muskoxen rumen – a metatranscriptomic approach. PLoS One 6, e20521.

Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto J-M, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Dore J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD and Wang J 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464, 59–65.

Quince C, Lanzen A, Curtis TP, Davenport RJ, Hall N, Head IM, Read LF and Sloan WT 2009. Accurate determination of microbial diversity from 454 pyrosequencing data. Nature Methods 6, 639–641.

Raes J and Bork P 2008. Molecular eco-systems biology: towards an understanding of community function. Nature Reviews Microbiology 6, 693–699.

Reeve JN, Nolling J, Morgan RM and Smith DR 1997. Methanogenesis: genes, genomes, and who's on first? Journal of Bacteriology 179, 5975–5986.

Rincon MT, Dassa B, Flint HJ, Travis AJ, Jindou S, Borovok I, Lamed R, Bayer EA, Henrissat B, Coutinho PM, Antonopoulos DA, Berg Miller ME and White BA 2010. Abundance and diversity of dockerin-containing proteins in the fiber-degrading rumen bacterium, *Ruminococcus flavefaciens* FD-1. PLoS One 5, e12476.

Rohwer F and Youle M 2011. Consider something viral in your research. Nature Reviews Microbiology 9, 308–309.

Rosenzweig RF, Sharp RR, Treves DS and Adams J 1994. Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. Genetics 137, 903–917.

Rosselló-Mora R and Amann R 2001. The species concept for prokaryotes. FEMS Microbiology Reviews 25, 39–67.

Russell JB, Muck RE and Weimer PJ 2009. Quantitative analysis of cellulose degradation and growth of cellulolytic bacteria in the rumen. FEMS Microbiology Ecology 67, 183–197.

Sadet S, Martin C, Meunier B and Morgavi DP 2007. PCR-DGGE analysis reveals a distinct diversity in the bacterial population attached to the rumen epithelium. Animal 1, 939–944.

Samuel BS, Hansen EE, Manchester JK, Coutinho PM, Henrissat B, Fulton R, Latreille P, Kim K, Wilson RK and Gordon JI 2007. Genomic and metabolic

adaptations of *Methanobrevibacter smithii* to the human gut. Proceedings of the National Academy of Sciences of the United States of America 104, 10643–10648.

Sanger F, Nicklen S and Coulson AR 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 74, 5463–5467.

Schadt EE, Turner S and Kasarskis A 2010. A window into third-generation sequencing. Human Molecular Genetics 19, R227–R240.

Shanks OC, Kelty CA, Archibeque S, Jenkins M, Newton RJ, McLellan SL, Huse SM and Sogin ML 2011. Community structures of fecal bacteria in cattle from different animal feeding operations. Applied and Environmental Microbiology 77, 2992–3001.

Shedova E, Berezina O, Lunina N, Zverlov V, Schwarz W and Velikodvorskaya G 2009. Cloning and characterisation of a large metagenomic DNA fragment containing glycosyl-hydrolase genes. Molecular Genetics, Microbiology and Virology 24, 12–16.

Simon C and Daniel R 2011. Metagenomic analyses: past and future trends. Applied and Environmental Microbiology 77, 1153–1161.

Stevens CE and Hume ID 1998. Contributions of microbes in vertebrate gastrointestinal tract to production and conservation of nutrients. Physiological Reviews 78, 393–427.

Stevenson DM and Weimer PJ 2007. Dominance of *Prevotella* and low abundance of classical ruminal bacterial species in the bovine rumen revealed by relative quantification real-time PCR. Applied Microbiology and Biotechnology 75, 165–174.

Suen G, Weimer PJ, Stevenson DM, Aylward FO, Boyum J, Deneke J, Drinkwater C, Ivanova NN, Mikhailova N, Chertkov O, Goodwin LA, Currie CR, Mead D and Brumm PJ 2011. The complete genome sequence of *Fibrobacter succinogenes* S85 reveals a cellulolytic and metabolic specialist. PLoS One e18814.

Sundset M, Edwards J, Cheng Y, Senosiain R, Fraile M, Northwood K, Præsteng K, Glad T, Mathiesen S and Wright A-D 2009. Molecular diversity of the rumen microbiome of Norwegian Reindeer on natural summer pasture. Microbial Ecology 57, 335–348.

Teather RM 2001. Community genomics – the key to the rumen? In Advances in beef cattle science (ed. KA Beauchemin and DHJ Crews), pp. 228–233. Livestock Sciences Section, Lethbridge Research Centre, Agriculture and Agri-Food Canada (AAFC), Lethbridge, Canada.

The Bovine Genome Sequencing Analysis Consortium, Elsik CG, Tellam RL, Worley KC, Gibbs RA, Muzny DM, Weinstock GM, Adelson DL, Eichler EE, Elnitski L, Guigo R, Hamernik DL, Kappes SM, Lewin HA, Lynn DJ, Nicholas FW, Reymond A, Rijnkels M, Skow LC, Zdobnov EM, Schook L, Womack J, Alioto T, Antonarakis SE, Astashyn A, Chapple CE, Chen H-C, Chrast J, Camara F, Ermolaeva O, Henrichsen CN, Hlavina W, Kapustin Y, Kiryutin B, Kitts P, Kokocinski F, Landrum M, Maglott D, Pruitt K, Sapojnikov V, Searle SM, Solovyev V, Souvorov A, Ucla C, Wyss C, Anzola JM, Gerlach D, Elhaik E, Graur D, Reese JT, Edgar RC, McEwan JC, Payne GM, Raison JM, Junier T, Kriventseva EV, Eyras E, Plass M, Donthu R, Larkin DM, Reecy J, Yang MQ, Chen L, Cheng Z, Chitko-McKown CG, Liu GE, Matukumalli LK, Song J, Zhu B, Bradley DG, Brinkman FSL, Lau LPL, Whiteside MD, Walker A, Wheeler TT, Casey T, German JB, Lemay DG, Maqbool NJ, Molenaar AJ, Seo S, Stothard P, Baldwin CL, Baxter R, Brinkmeyer-Langford CL, Brown WC, Childers CP, Connelley T, Ellis SA, Fritz K, Glass EJ, Herzig CTA, Iivanainen A, Lahmers KK, Bennett AK, Dickens CM, Gilbert JGR, Hagen DE, Salih H, Aerts J, Caetano AR, Dalrymple B, Garcia JF, Gill CA, Hiendleder SG, Memili E, Spurlock D, Williams JL, Alexander L, Brownstein MJ, Guan L, Holt RA, Jones SJM, Marra MA, Moore R, Moore SS, Roberts A, Taniguchi M, Waterman RC, Chacko J, Chandrabose MM, Cree A, Dao MD, Dinh HH, Gabisi RA, Hines S, Hume J, Jhangiani SN, Joshi V, Kovar CL, Lewis LR, Liu Y-s, Lopez J, Morgan MB, Nguyen NB, Okwuonu GO, Ruiz SJ, Santibanez J, Wright RA, Buhay C, Ding Y, Dugan-Rocha S, Herdandez J, Holder M, Sabo A, Egan A, Goodell J, Wilczek-Boney K, Fowler GR, Hitchens ME, Lozado RJ, Moen C, Steffen D, Warren JT, Zhang J, Chiu R, Schein JE, Durbin KJ, Havlak P, Jiang H, Liu Y, Qin X, Ren Y, Shen Y, Song H, Bell SN, Davis C, Johnson AJ, Lee S, Nazareth LV, Patel BM, Pu L-L, Vattathil S, Williams Jr RL, Curry S, Hamilton C, Sodergren E, Wheeler DA, Barris W, Bennett GL, Eggen A, Green RD, Harhay GP, Hobbs M, Jann O, Keele JW, Kent MP, Lien S, McKay SD, McWilliam S, Ratnakumar A, Schnabel RD, Smith T, Snelling WM, Sonstegard TS, Stone RT, Sugimoto Y, Takasuga A, Taylor JF, Van Tassell CP, MacNeil MD, Abatepaulo ARR, Abbey CA, Ahola V, Almeida IG, Amadio AF, Anatriello E, Bahadue SM, Biase FH, Boldt CR, Carroll JA, Carvalho WA, Cervelatti EP, Chacko E, Chapin JE, Cheng Y, Choi J, Colley AJ, de Campos TA, De Donato M, Santos IKFdM, de Oliveira CJF, Deobald H, Devinoy E, Donohue KE, Dovc P, Eberlein A, Fitzsimmons CJ, Franzin AM, Garcia GR, Genini S, Gladney CJ, Grant JR, Greaser ML, Green JA, Hadsell DL, Hakimov HA, Halgren R, Harrow JL, Hart EA, Hastings N, Hernandez M, Hu Z-L, Ingham A, Iso-Touru T, Jamis C, Jensen K, Kapetis D, Kerr T, Khalil SS, Khatib H, Kolbehdari D, Kumar CG, Kumar D, Leach R, Lee JCM, Li C, Logan KM, Malinverni R, Marques E, Martin WF, Martins NF, Maruyama SR, Mazza R, McLean KL, Medrano JF, Moreno BT, More DD, Muntean CT, Nandakumar HP, Nogueira MFG, Olsaker I, Pant SD, Panzitta F, Pastor RCP, Poli MA, Poslusny N, Rachagani S, Ranganathan S, Razpet A, Riggs PK, Rincon G, Rodriguez-Osorio N, Rodriguez-Zas SL, Romero NE, Rosenwald A, Sando L, Schmutz SM, Shen L, Sherman L, Southey BR, Lutzow YS, Sweedler JV, Tammen I, Telugu BPVL, Urbanski JM, Utsunomiya YT, Verschoor CP, Waardenberg AJ, Wang Z, Ward R, Weikard R, Welsh Jr TH, White SN, Wilming LG, Wunderlich KR, Yang J and Zhao F-Q 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science 324, 522–528.

The Human Microbiome Jumpstart Reference Strains C 2010. A catalog of reference genomes from the human microbiome. Science 328, 994–999.

The International Sheep Genomics Consortium, Archibald AL, Cockett NE, Dalrymple BP, Faraut T, Kijas JW, Maddox JF, McEwan JC, Hutton Oddy V, Raadsma HW, Wade C, Wang J, Wang W and Xun X 2010. The sheep genome reference sequence: a work in progress. Animal Genetics 41, 449–453.

Tringe SG and Hugenholtz P 2008. A renaissance for the pioneering 16S rRNA gene. Current Opinion in Microbiology 11, 442–446.

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R and Gordon JI 2007. The human microbiome project. Nature 449, 804–810.

von Mering C, Hugenholtz P, Raes J, Tringe SG, Doerks T, Jensen LJ, Ward N and Bork P 2007. Quantitative phylogenetic assessment of microbial communities in diverse environments. Science 315, 1126–1130.

Wallace RJ 2008. Gut microbiology – broad genetic diversity, yet specific metabolic niches. Animal 2, 661–668.

Wang FC, Li F, Chen GJ and Liu WF 2009. Isolation and characterization of novel cellulase genes from uncultured microorganisms in different environmental niches. Microbiological Research 164, 650–657.

Warnecke F, Luginbuhl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, Cayouette M, McHardy AC, Djordjevic G, Aboushadi N, Sorek R, Tringe SG, Podar M, Martin HG, Kunin V, Dalevi D, Madejska J, Kirton E, Platt D, Szeto E, Salamov A, Barry K, Mikhailova N, Kyrpides NC, Matson EG, Ottesen EA, Zhang XN, Hernandez M, Murillo C, Acosta LG, Rigoutsos I, Tamayo G, Green BD, Chang C, Rubin EM, Mathur EJ, Robertson DE, Hugenholtz P and Leadbetter JR 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature 450, U560–U517.

Weber M, Teeling H, Huang S, Waldmann J, Kassabgy M, Fuchs BM, Klindworth A, Klockow C, Wichels A, Gerdts G, Amann R and Glockner FO 2011. Practical application of self-organizing maps to interrelate biodiversity and functional data in NGS-based metagenomics. ISME Journal 5, 918–928.

Weimer PJ 1998. Manipulating ruminal fermentation: a microbial ecological perspective. Journal of Animal Science 76, 3114–3122.

Weimer PJ and Odt CL 1995. Cellulose degradation by ruminal microbes: physiological and hydrolytic diversity among ruminal cellulolytic bacteria. In Enzymatic degradation of insoluble carbohydrates (ed. J Saddler and M Penner), pp. 291–304. American Chemical Society, Washington, DC.

Weimer PJ, Russell JB and Muck RE 2009. Lessons from the cow: what the ruminant animal can teach us about consolidated bioprocessing of cellulosic biomass. Bioresource Technology 100, 5323–5331.

Weiss WP, Wyatt DJ and McKelvey TR 2008. Effect of feeding *Propionibacteria* on milk production by early lactation dairy cows. Journal of Dairy Science 91, 646–652.

Welkie DG, Stevenson DM and Weimer PJ 2010. ARISA analysis of ruminal bacterial community dynamics in lactating dairy cows during the feeding cycle. Anaerobe 16, 94–100.

Whitford MF, Forster RJ, Beard CE, Gong J and Teather RM 1998. Phylogenetic analysis of rumen bacteria by comparative sequence analysis of cloned 16S rRNA genes. Anaerobe 4, 153–163.

Williams AG and Coleman GS 1992. The Rumen Protozoa. Springer-Verlag New York Inc., New York.

Wilson DB 2011. Microbial diversity of cellulose hydrolysis. Current Opinion in Microbiology 14, 259–263.

Wolin M, Miller T and Stewart C 1997. Microbe–microbe interactions. In The rumen microbial ecosystem (ed. P Hobson and C Stewart), pp. 467–491. Chapman & Hall, London.

Wong D, Chan VJ, McCormack AA and Batt SB 2010a. Cloning and characterization of an exo-xylogucanase from rumenal microbial metagenome. Protein and Peptide Letters 17, 803–808.

Wong D, Chan VJ, McCormack AA and Batt SB 2010b. A novel xyloglucan-specific endo-β-1,4-glucanase: biochemical properties and inhibition studies. Applied Microbiology and Biotechnology 86, 1463–1471.

Wong DWS, Chan VJ and Batt SB 2008. Cloning and characterization of a novel exo-α-1,5-L-arabinanase gene and the enzyme. Applied Microbiology and Biotechnology 79, 941–949.

Wong DWS, Chan VJ and McCormack AA 2009. Functional cloning and expression of a novel endo-α-1,5-L-arabinanase from a metagenomic library. Protein and Peptide Letters 16, 1435–1441.

World Bank 2008. Rising Food and Fuel Prices: Addressing the Risks to Future Generations. Retrieved September 1, 2009, from http://siteresources.worldbank.org/DEVCOMMEXT/Resources/Food-Fuel.pdf

Xu J and Gordon JI 2003. Inaugural article: honor thy symbionts. Proceedings of the National Academy of Sciences of the United States of America 100, 10452–10459.

Yanez-Ruiz DR, Hart KJ, Martin-Garcia AI, Ramos S and Newbold CJ 2008. Diet composition at weaning affects the rumen microbial population and methane emissions by lambs. Australian Journal of Experimental Agriculture 48, 186–188.

Yáñez-Ruiz DR, Macías B, Pinloche E and Newbold CJ 2010. The persistence of bacterial and methanogenic archaeal communities residing in the rumen of young lambs. FEMS Microbiology Ecology 72, 272–278.

Zhao S, Wang J, Bu D, Liu K, Zhu Y, Dong Z and Yu Z 2010. Novel glycoside hydrolases identified by screening a Chinese Holstein dairy cow rumen-derived metagenome library. Applied and Environmental Microbiology 76, 6701–6705.

Zhou M, Hernandez-Sanabria E and Guan LL 2010. Characterization of variation in rumen methanogenic communities under different dietary and host feed efficiency conditions, as determined by PCR-denaturing gradient gel electrophoresis analysis. Applied and Environmental Microbiology 76, 3776–3786.

Zhu YX, Wang JQ, Ma RL, Huang L and Dong ZY 2007. Construction and analysis of rumen bacterial artificial chromosome library from a dairy cow rumen microflora. Wei Sheng Wu Xue Bao (Acta Microbiologica Sinica) 47, 213–216.