Metagenomics and its Application in Rumen Ecosystem: Potential Biotechnological Prospects

Shakira Ghazanfar1,2 and Atiya Azim1
1Department of Animal Nutrition, Animal Sciences Institute, National Agricultural Research Centre, Islamabad, Pakistan
2Department of Animal Genomics and Biotechnology, National University of Agricultural Sciences, Islamabad, Pakistan

Abstract: Genetic and biological diversity of microorganisms is an important area of scientific research. Considering the importance of ruminants in livestock strategies their ability to convert locally available feedstuff to animal products should be improved. Recent advances in the molecular biology and genomics now offer new opportunities to conduct a more holistic examination of the structure and function of rumen communities. To understand the complex microbial communities function and how microbes interact within their niches represents a major challenge for rumen microbiologists today. Metagenomics has the potential for providing insight into the functional dimensions of the rumen genomic database and will help to achieve a major goal of rumen microbiology; the complexities of microbial communities function and interaction among these microbes. This review summarizes the molecular methods of culture-independent insight ‘metagenomics’ and their recent application to studies of the rumen ecosystem for enhancing the livestock productivity.

Key words: Metagenomics, rumen ecosystem, biotechnology

INTRODUCTION
The microbial populations are vital to life on the earth and are of enormous practical significance in medicine; engineering and agriculture (Sloan et al., 2006). Microbes have many of the properties similar to complex organism like humans. They also exhibit unique properties such as ability to degrade waste products. As a result, the genetic and biological diversity of the microorganisms is an important area of scientific research (Singh et al., 2008). The global microbial diversity presents an enormous, largely untapped genetic and biological pool that could be exploited for the recovery of novel genes, biomolecules for metabolic pathways and various valuable products (Cowan, 2000). But unfortunately, Current research indicated that more than 99% of the microbes in the environment can not be readily cultivated (Harada, 2003; Kamagata and Tamaki, 2005; Rappe and Giovannoni, 2003; Sekiguchi, 2006). Thus, most microbes not been described and accessed for biotechnology or any basic research. In fact, most of the microbial species in many environments have been never described and this situation will not change until new cultural technologies are devolved. Additionally, many approaches currently used to explore the diversity and potential of microbial communities are biased because of the limitations of cultivation methods. To overcome the difficulties and limitations associated with cultivation techniques, several DNA-based molecular methods have been developed. In general, methods based on 16S rRNA gene analysis provide extensive information about the taxa and species present in an environment. However, these data usually provide only little if any information about the functional role of the different microbes within the community and the genetic information they contain of microbial niches (Streit and Schmitz, 2004).

Metagenomics is a rapidly growing field of research that aims at studying uncultured organisms to understand the true diversity of microbes, their functions, cooperation and evolution, in environments such as soil, water, ancient remains of animals, or the digestive system of animals and humans (Huson et al., 2009). This review summarizes the molecular methods of culture-independent insight ‘metagenomics’ and their recent application to studies of the rumen ecosystem for enhancing the livestock productivity.

What is metagenomics: The term “metagenomics” was first coined by Handelsman et al. (1998) to study the genomes from all microbes in a particular environment as opposed to the genome from one organism isolated from the environment and cultured in vitro. In principle, any study that addresses all the individuals in a microbial community as a single genomic pool can be seen as an exercise in metagenomics (Kowaichuk et al., 2007). Initially, noncultured microflora and ancient DNA investigations had been the prime targets of metagenomic studies, but presently the technology has been applied to study an array of microbial diversities like deep-sea aquatic microflora, soil microbes and GI

1309
ecosystem of human and animals (Lu et al., 2007; Shanks et al., 2006). Studies have revealed that only 0.001-0.1% of the total microbes in sea water, 0.25% in freshwater, 0.25% in sediments and only 0.3% soil microorganisms could be cultivable in vitro (Amann et al., 1995).

The current metagenomic studies have largely progressed due to the construction of efficient gene cloning vectors like bacterial artificial chromosomes (BACs) or cosmids, etc (Babcock et al., 2007; Xu, 2008) which allow cloning and expression of larger and complex DNA segments or genes and the development of methods for generation and analysis of the data.

Metagenome technologies, DNA extraction, library construction, screening: Metagenome analyses are usually initiated by the isolation of environmental DNAs. A major difficulty associated with the metagenome approach is related to the contamination of purified DNA with polyphenolic compounds, which are copurified with the DNA. These compounds are difficult to remove and it is well known that polyphenols also interfere with enzymatic modifications of the isolated DNA (Tsai and Olson, 1992). In addition, for marine samples often large amounts of water have to be filtered to obtain sufficient DNA for cloning (Beja et al., 2000). As a result, the construction of environmentally derived DNA libraries with large inserts is often hindered because of the poor quality of the isolated DNA. These known difficulties associated with the construction of libraries directly derived from environmental DNA samples have forced many laboratories to isolate DNA from the metagenome of a microbial community after pre-cultivation in the laboratory. Although it is expected that laboratory enrichment cultures bear only a limited biodiversity this technique has proven to be highly efficient for the rapid isolation of large DNA fragments and for cloning of operons and genes with high biotechnological value (Vogel et al., 2003; Knietsch et al., 2003; Entcheva et al., 2001; Healy et al., 1995). DNA isolation and purification is followed by the construction of DNA libraries in suitable cloning vectors and host strains. The classical approach includes the construction of small insert libraries (<10 kb) in a standard sequencing vector and in Escherichia coli as a host strain (Henne et al., 1999). However, small insert libraries do not allow detection of large gene clusters or operons. To circumvent this limitation researchers have been employing large insert libraries, such as cosmid DNA libraries (mostly in the pWE15 vector of Stratagene) with insert sizes ranging from 25-35 kb (Entcheva et al., 2001) and/or Bacterial Artificial Chromosome (BAC) libraries with insert sizes up to almost 200 kb (Beja et al., 2000; Rondon et al., 2000). Additionally, the construction of fosmid libraries with inserts of 40 kb of foreign DNA has been reported (Beja et al., 2002). E. coli is still the preferred host for the cloning and expression of any metagenome-derived genes. Only very recently have other hosts such as Streptomyces lividans been employed to identify genes involved in the biosynthesis of novel antibiotics (Courtois et al., 2003). Also, to our knowledge metagenomic libraries are currently developed in other Gram-negative hosts by several laboratories working in the field, which will become available soon.

The development of these new host strains will increase screening efficiency tremendously. Functional searches for novel genes in metagenomic libraries have often been performed using highly sophisticated picking and pipetting robots. In many of the recent publications large clone-libraries have been screened. Often several hundred thousand clones have been analyzed to detect less than ten active clones in a single screen (Henne et al., 1999; Henne et al., 2000; Majernik et al., 2001). This is mainly owing to a lack of efficient transcription of the metagenome-derived genes in the host strain. This effect might be worsened by a weak translation in combination with a poor secretion of the foreign protein by the employed host strain. Also, in many cases it can be expected that the desired protein is not folded correctly because required chaperones are absent in the host strain. Furthermore, cofactors might not be synthesized by the foreign host strain and/or not inserted correctly into the recombinant metagenomic protein. Finally, a different codon usage could be a reason for poor protein expression and low activities. Several laboratories are currently working on solutions for some of these problems and constructing novel vectors and strains. Other avenues to pursue include the development of highly sensitive screening methods using fluorogenic substrates. Large-scale sequencing projects such as the one initiated by Craig Venter for the metagenome of the Sargasso Sea resulted in the identification of numerous novel genes and is a very famous example of sequence-based metagenome analyses (Venter et al., 2004). Similar approaches have been initiated by European laboratories to sequence complete or partial metagenomes of a phylogenetically highly diverse biofilm (Schmeisser et al., 2003). Complementary to this, Tyson and coworkers described the nearly complete sequence analyses of the metagenome of an acidophilic biofilm, which was characterized by a low biodiversity (Tyson et al., 2004) A completely novel strategy has been developed by searching through metagenome DNA sequences and by using microarray technology (Sebat et al., 2003). The microarray profiling offers an effective approach for rapidly identifying and characterizing many clones. It could potentially be used for the isolation of large numbers of conserved genes. Similarly, other laboratories have been using degenerated PCR primers for the isolation of metagenome-derived genes. Among these are genes involved in the degradation of a-
halocarboxylic acids (Marchesi and Weightman, 2003) and genes encoding for novel hydrolases (Bell et al., 2002).

Rumen function and metagenomics: Ruminants will probably be important in livestock strategies to assist the poor (Delgado, 2005), therefore their ability to convert locally available feedstuffs to animal products should be improved. It’s well studied that microbial community inhabiting in the rumen is characterized by its high population density, wide diversity and interactive complexity (Duan et al., 2006). This microbial community is responsible for the bioconversion of lignocellulosic feeds into volatile fatty acids (Kamra, 2005). The goal of rumen biotechnologists is to manipulate the ruminal microbial ecosystem to improve the efficiency of feed (Khampa and Wanapat, 2007). It is well known that biotechnology has a continuous demand for novel genes and enzymes and compounds (Christel et al., 2007). The rumen microbial diversity represents a vast genetic bounty that may be exploited for the discovery of novel genes, entire metabolic pathways and potentially valuable end-products thereof (Frank and Pace, 2008), but the successful development for the discovery of some novel genes or microbes have not yet been achieved.

For the successful development, it is important to study and to intend more basic characteristics of rumen microbes in the ecosystem. Otherwise future development of rumen and ruminants studies would not be expecting (Ling, 1994). Rumen microbiology has made a significant contribution to the understanding of ruminant nutrition. However, further progress in research has been hindered by the incomplete analysis of the rumen microbiota comprised of bacteria, protozoa and fungi, most of which remain uncharacterized due to the difficulties in isolation and cultivation (Yasuo, 2006). Rumen microbial populations could be under-estimated by adopting the traditional techniques (Weidong et al., 2008). Traditional methods for culturing microorganisms fail to represent the scope of microbial diversity in nature and they limit analysis to those that grow under laboratory condition (Hugenholtz, 2002; Rondon et al., 2000).

Over the past decade, extensive suites of molecular based approaches have been developed that have enabled the study of uncultured microorganisms. But the progress in exploiting the rumen ecosystem is slow because until recently knowledge of rumen microbiology is primarily based on classical culture based techniques (isolation, enumeration and nutritional characterization) which probably only account for 10-20% of the rumen microbial population (Duan et al., 2006). When Hungate (1966) published The Rumen and Its Microbes, about 23 bacterial species were thought to play prominent roles in ruminal metabolism; by 1996, the number exceeded 200 (Krause and Russell, 1998). When several discrete ribosomal DNA libraries were analyzed, 341 operational taxonomic units of organisms were identified (Edwards et al., 2004), this indicated that culture-based estimates of ruminal organisms greatly underestimated ruminal diversity. Large population of rumen microbiogism is non cultural but is active in rumen fermentation (Kamra, 2005) and it has been well recognized that there is a vast amount of information held within genomes of uncultured microorganisms (Handelsman, 2004). Therefore it is essential to search for some better technique that focused on identifying the microbes and microbial mechanisms of deriving the nutrients from low quality forages, enhanced dietary fiber digestion by the selected elite rumen flora and studying the nutrients-host tissue interactions (Singh et al., 2009). Recent advances in molecular biology and genomics now offer new opportunities to content a more holistic examination of the rumen microbial communities, including those that cannot readily cultural (Firkins et al., 2007).

Application of rumen metagenomics: The development and application of metagenomics has allowed access to the uncultivated ecosystem and insight into metabolic capabilities as yet uncultured microbial communities. Some examples of application of metagenomics are as following:

Genomic sequences database of the rumen microbes: Sequencing the genomes of individual rumen microbes and determining the function of their encoded genes promises to transform our understanding of the microbiology of the rumen (Attwood et al., 2008). (Table 1) list the ruminal bacteria, that at the time of writing are the subjects of genome sequencing projects as well as the sequencing genome that are available for strain and the website. Although the total number of sequencing rumen microbial genome is comparably low, there is sequence data available for phylogenetically related specie and strain.

This Genomic sequences database support the examination of genome composition and organization and opportunity to extant our understanding of rumen microbes beyond the degradative and metabolic characteristic predicatively relevant to the host ruminant nutrition.

Identification of novel enzymes/microbes from rumen: The metagenomics need to be exploited to screen and identify novel microbes and biomolecules from the GI tract of the livestock ruminants adapted to the forages or diets enriched with high fiber and an array of antinutritional Plant Secondary Metabolites (PSMs) such as tannin-polyphenols. Feral herbivores or the migratory livestock species like goats and sheep could harbor a wealth of valuable GI microorganisms that could be
developed and then used as probiotics or direct fed microbials (to enhance rumen productivity) and sources of various hydrolytic enzymes for promoting livestock nutrition, health and industrial development (Singh et al., 2008) (Table 2).

A metagenome expression library of bulk DNA extracted from the rumen content of a dairy cow was established in a phage lambda vector and activity-based screening employed to explore the functional diversity of the microbial flora. Twenty-two clones specifying distinct hydrolytic activities (12 esterases, nine endo-β-1, 4-glucanases and one cyclodextrinase) were identified in the library and characterized. Sequence analysis of the retrieved enzymes revealed that eight (36%) were entirely new and formed deep-branched phylogenetic lineages with no close relatives among known esterases and glycosyl-hydrolases (Ferrer et al., 2007). A novel enzyme, RA-04, belonging to the α-amylase family was obtained after expression of metagenomic DNA from rumen fluid. The studies have demonstrated the usefulness of a metagenomic approach to gain novel debranching enzymes, important for the bread/food industries, from microbial environments with a high rate of plant polymer turnover, exemplified by the cow rumen (Ferrer et al., 2007; Ferrer et al., 2005).

In another study, RL5 (EMBL/DBJ/GenBankTM accession number AM269758 [GenBank]), a gene coding for a novel Laccases (polyphenol oxidase) was identified through activity screening of a metagenome expression library from the bovine rumen micro flora. Laccases in the rumen may play an important role in ryegrass lignin digestion so the RL5 enzyme may have biotechnological potential for exploitation in pasture-fed animals and in pasture grasses. Finally, the study demonstrates the power and utility of activity-based metagenomics for the exploration of functional diversity space and the discovery of novel enzymes with laccase activity in a protein lacking homology to any described polyphenol oxidase (Belouqui et al., 2006).

A unique multifunctional glycosyl hydrolase was discovered by screening as environmental DNA library prepared from a microbial consortium collected from cow rumen. This enzyme is active on several different, linked substrate and process manner xylan activesa (Nisha et al., 2007).

Identification of uncultured methanogens: Methanogens are member of the domain Archaea, fall within the kingdom euarchaeota (Woese et al., 1990). They are obligate anaerobes and produced methane as a major catabolic product (Bergey, 1994). Interest in methanogens from ruminants has resulted from the role of methane from the fact that cattle lose 6% of ingestion energy as methane (Johnson and Johnson, 1995). Methanogens in the rumen have been identified by different molecular techniques. A Temporal Temperature Gradient Gel Electrophoresis (TGGE) method developed to determine the diversity of methanogens in cattle and sheep rumens showed that uncultured methanogens account for the majority of methanogenic archaea in the rumen (Nicholson et al., 2007). Denaturing Gradient Gel Electrophoresis (DGGE) enables separation of double-stranded DNA fragments up to 500bp in length utilizing either denaturing or temperature gradient gel (Kocherginskaya et al., 2005). Now a days Denaturing Gradient Gel Electrophoresis (DGGE) would be used for the identification of uncultured methanogens. Elaborating the molecular mechanisms of association patterns between archaea and rumen protozoa would be helpful in developing strategies to reduce methane emissions by dietary or genetic manipulation of the rumen ecosystem (Singh et al., 2008).

Differentiating and quantitative determination rumen biomass: Another important application of microbial metagenomics in animal nutrition is the quantitative determination of total rumen microbial biomass and differentiating the bacterial and protozoal biomass. Rapid profiling procedures, such as Denaturing Gradient Gel Electrophoresis (DGGE), can be used to infer likely differences in community structure of bacteria and archaea among numerous replicates of animals.

### Table 1: Genome Sequence projects for rumen microorganisms. The institution coordinating the projects and websites for access to date data are also shown.

<table>
<thead>
<tr>
<th>Rumen microbes</th>
<th>Stain and coverage</th>
<th>Institution</th>
<th>websites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibrobacter succinogenes</td>
<td>Strain S 85 closed</td>
<td>TIGR</td>
<td><a href="http://www.tigr.org/db/rumenomics">www.tigr.org/db/rumenomics</a></td>
</tr>
<tr>
<td>Ruminococcus albus</td>
<td>Strain 8 closed</td>
<td>TIGR</td>
<td><a href="http://www.tigr.org/db/rumenomics">www.tigr.org/db/rumenomics</a></td>
</tr>
<tr>
<td>Prevotella ruminicola</td>
<td>Strain 23 closed</td>
<td>TIGR</td>
<td><a href="http://www.tigr.org/db/rumenomics">www.tigr.org/db/rumenomics</a></td>
</tr>
<tr>
<td>Prevotella bryantii</td>
<td>Strain B. 4, Draft (8x)</td>
<td>TIGR</td>
<td><a href="http://www.tigr.org/db/rumenomics">www.tigr.org/db/rumenomics</a></td>
</tr>
<tr>
<td>Ruminococcus flavefaciens</td>
<td>Strain FD -1 Draft (2x)</td>
<td>University of Illinois</td>
<td><a href="http://www.biotech.uarc.edu">www.biotech.uarc.edu</a></td>
</tr>
</tbody>
</table>

### Table 2: Some rumen microbes and enzymes identified using metagenomic tools.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Enzymes/microbial studies</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferrer et al., 2005</td>
<td>Novel enzyme RA.04 belonging to the alpha-amylose family</td>
<td>Bovine rumen</td>
</tr>
<tr>
<td>Belouqui et al., 2006</td>
<td>Novel gene RI-5 encoding polyphenol oxidase</td>
<td>Bovine rumen</td>
</tr>
<tr>
<td>Matthew et al., 2007</td>
<td>Novel methanogens</td>
<td>Cattle, Sheep rumen</td>
</tr>
<tr>
<td>Lammle et al., 2007</td>
<td>Rumen microbes</td>
<td>Bovine rumen (SSH)</td>
</tr>
<tr>
<td>Palackal et al., 2007</td>
<td>Multifunctional glycosyl hydrolase</td>
<td>Bovine rumen</td>
</tr>
</tbody>
</table>

---

and times after feeding diets that are more representative of intense ruminant animal production (Firkins et al., 2007).

Ruminal nitrogen metabolism: A better understanding of mechanistic process altering the production and uptake of amino nitrogen will help the livestock nutritionists to improve the overall conversion of dietary nitrogen into microbial protein. It will provide key information needed to further improve the mechanistic models describing the rumen function and evaluating dietary conditions that influence the efficiency of conversion of dietary nitrogen into milk protein (Firkins et al., 2007).

Conclusion: This review is intended to focus the reader's attention on the potential of exploiting metagenomics, specifically the use of metagenomic libraries constructed from unique rumen ecosystem as an approach to successfully exploit the largely "untapped" resources within various rumen ecosystems. The microbial community inhabiting in the rumen is characterized by its high population density, wide diversity and interactive complexity. These diverse ecosystems are potentially very useful sources for novel enzymes with unique properties and great biotechnological potential. Only a small proportion of the bioresources of rumen microbiota have thus far been examined and an even smaller proportion has been exploited. Given our present inability to cultivate the vast majority of microbes from this ecosystem the metagenomic approaches outlined in this review offer the only methodology currently available to access these unique and useful bioresources. There is an ongoing need for a wide range of novel genes and enzymes which are required to improve fiber digestion, increasing digestibility of low quality forage, by the selected elite rumen flora and studying the nutrients-host tissue interaction. Rumen metagenomics coupled with biotechnology has the potential to contribute to all these pressing needs. These technologies have the potential to revolutionize the understanding of rumen function and will overcome the limitations of classical based techniques, including isolation and taxonomic identification of strains important to efficient rumen function and better understanding of the roles of microorganisms in relation to achieving high productivity and decreasing environmental pollutants.

REFERENCES


