METAGENOMICS AND ITS APPLICATION IN RUMEN ECOSYSTEM: POTENTIAL BIOTECHNOLOGICAL PROSPECTS

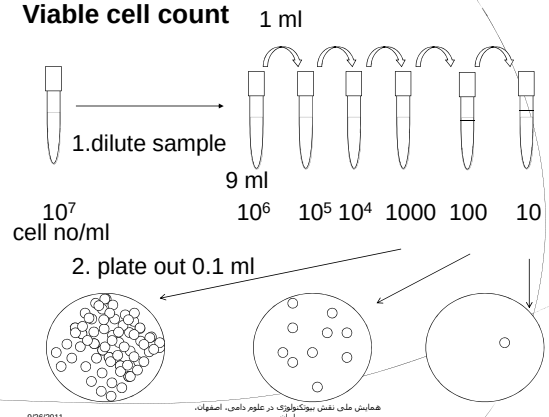
Introduction

- 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).

Microbial ecology seeks to answer 3 questions:

1. Who's there?
2. Who's active?
3. What are they doing?

Viable cell count



Molecular microbial ecology seeks to answer 3 questions:

1. Who's there?

2. Who's active?

3. What are they doing?

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.*



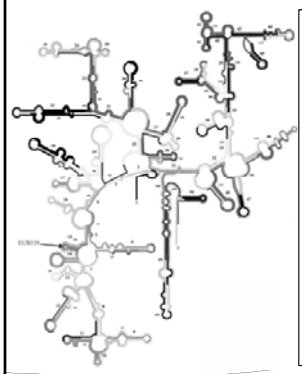
Molecular methods for enumeration

- Genetic-based methods are rapidly replacing conventional detection and enumeration methods in microbiology.
- These methods are mainly based on small subunit ribosomal RNA sequences.

Why molecular methods?

- The majority of environmental microbes are not easily cultured by usual lab methods.
- Microbes are morphologically indistinct.
- Culture-independent molecular methods can be used to distinguish bacterial populations and to describe microbial communities.
- Most of these methods involve PCR amplification of DNA extracted from natural samples.

The 16S ribosomal RNA molecule



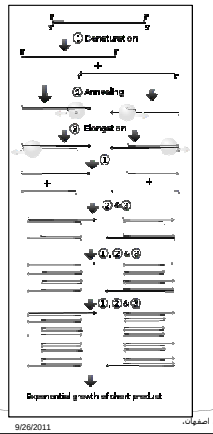
The 16S rRNA molecule is a major component of the small ribosomal subunit. It has approximately 1500 ribonucleotides. This single-stranded rRNA molecule has an intricate secondary structure with extensive intrachain base pairing. The 16S rRNA forms a part of the ribosomal structure that is the site of protein biosynthesis resulting in the translation of messenger RNA. The 3' end of the bacterial 16S rRNA base-pairs with the Shine- Dalgarno sequence located upstream of the AUG initiation codon in mRNA during the initiation step of the translation process. This allows the mRNA to position itself on the ribosome. There is also evidence that 16S rRNA is directly involved in the interactions between the large and small ribosomal subunits.

By Dr. G.E. Fox,
Dept of Biology and Biochemistry
University of Houston

DNA extraction

Three steps:

1. **Lysis of the bacterial cells.**
   - mechanical means (physically breaking open the cells)
   - chemical means (using enzymes that break down the bacterial cell walls).
2. **Protecting the DNA from degradation.** Proteins are precipitated, leaving the DNA in solution. This removes enzymes (proteins) that can break down DNA.
3. **Purifying the DNA.** DNA is precipitated with ethanol, and cellular components that do not precipitate with ethanol are washed away. The precipitated DNA is dried and resuspended in water or buffer.



PCR – Polymerase chain reaction

Schematic drawing of the PCR cycle
(1) Denaturing at 94-96°C.
(2) Annealing at ~65°C
(3) Elongation at 72°C.
Four cycles are shown here. The blue lines represent the DNA template to which primers (red arrows) anneal that are extended by the DNA polymerase (light green circles), to give shorter DNA products (green lines), which themselves are used as templates as PCR progresses.

Source: Wikipedia PCR

Clone Libraries

- Isolate environmental DNA
- Amplify 16S rDNA using PCR
- Clone PCR product and sequence
- Compare to 16S rRNA database to identify sequence (closest relative)

***Bacillus* sp. ICPS6 16S rDNA**

```
      1 agtttgatcc tggctcagga cgaacgctgg cggcgtgcct aatacatgca agtcgagcgg
     61 acgggcgggg agcttgcttc cgcttggtta gcggcggacg ggtgagtaac acgtgggtaa
    121 cctgccctga agaccggaat aactccggga aaccggggct aataccggat aacaccgaag
    181 accgcatggt cttcggttga aaggcggctt cggctgccac ttacggatgg gcccgcggcg
    241 cattagctag ttggtgaggt aacggctcac caaggcgacg atgcgtagcc ggcctgagag
    301 ggtgacggc cacactggga ctgaaacacg gcccagactc ctacgggagg cagcagtagg
    361 gaatcttccg caatggacga aagtctgacg gagcgacgcc gcgtgagcga agaaggtctt
    421 cggatcgtaa agctctgttg ttaggaaga agaagtgccg ttcgaaacgg gcggcggtgt
    481 gacggtacct aacgagaaag cccggctaa ctacgtgcca gcagccgcgg taatacgtag
    541 gggcgcagcg ttgtccggaa ttattgggcg taaagcgcgc gcaggcggtc ccttaagtct
    601 gatgtgaaag cccacggctt aacgtggag ggtcattgga aactgggga cttgagtgca
    661 gaagaggaga gcggaattcc acgtgtagcg gtgaaatgcg tagagatgtg gaggaacacc
    721 agtggcgaag gcggctctgt ggtctgtaac tgacgctgag gcgcgaaagc gtggggagca
    781 aacaggatta gataccctgg tagtccacgc cgtaaacgat gagtgctaag tgttagaggg
    841 gttaaccctt tagtgctgta gctaacgcgt taagcactcc gcctggggag tacggccgca
    901 aggctgaaac tcaaaggaat tgacgggggc ccgcacaagc ggtggagcat gtggtttaat
    961 tcgaagcaac gcgaagaacc ttacctggtc ttgacatccc ctgacaaccc tggagacagg
   1021 gcgttcctcc cttgcgggag gacagggtga cagtggtgc atggttgtcg tcagctcgtg
   1081 tcgtgagatg ttgggttaag tcccgcaacg agcgcaaccc tcgccctag ttgccagcat
   1141 tcagttgggc actctagggg gactgccggc taaaagtcgg aggaaggtgg ggatgacgtc
   1201 aaatcatcat gccccttatg acctgggcta cacacgtgct acaatggcg gtacaaaggg
   1261 ctgcgaaccc gcgagggga gcgaatccca aaaagccgct ctcagttcgg attgcaggct
   1321 gcaactcgcc tgcatgaagc cggaatcgct agtaatcgcg gatcagcatg ccgcggtgaa
   1381 tacgttcccg ggccttgtac acaccgcccg tcacaccacg agagcttgca acacccgaag
   1441 tcggtgaggt aacccgaag ggagcagcc gcgaaggtg gggcaagtga ttggggtgaa
   1501 gtcgtaacaa
```


Phylogenetic tree

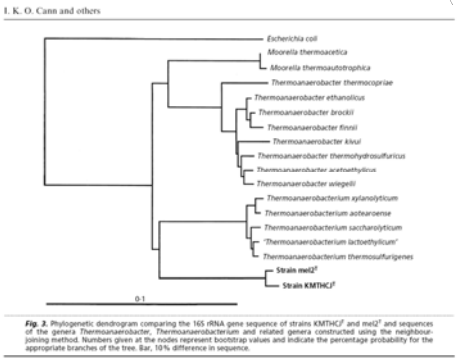


Fig. 3. Phylogenetic dendrogram comparing the 16S rRNA gene sequence of strains KMTHCJ and mel2 and sequences of the genera Thermoanaerobacter, Thermoanaerobacterium and related genera constructed using the neighbour-joining method. Numbers given at the nodes represent bootstrap values and indicate the percentage probability for the appropriate branches of the tree. Bar, 10% difference in sequence.

Rumen function and metagenomics:

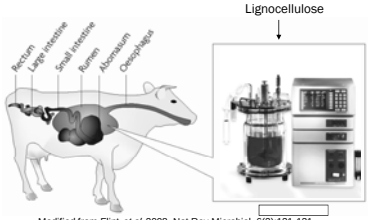
- Ruminants will probably be important in livestock strategies to assist the poor (Delgado, 2005), therefore their ability to convert locally available feedstuffs to animal products should be improved.
- It's well studied that microbial community inhabiting in the rumen is characterized by its high population density, wide diversity and interactive complexity (Duan *et al.*, 2006).

Rumen function and metagenomics:

- This microbial community is responsible for the bioconversion of lignocellulosic feeds into volatile **fatty acid**s (Kamra, 2005).
- The goal of rumen biotechnologists is to manipulate the ruminal microbial ecosystem to improve the efficiency of feed (Khampa and Wanapat, 2007).

The rumen is a highly efficient bioreactor

Substrates
Lignocellulose

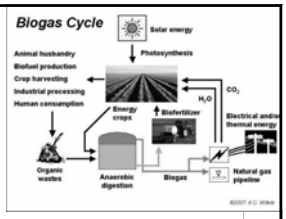


Fermentation Products
Fatty Acids
- acetate
- propionate
- butyrate

Gas
- methane
- CO_2

Modified from Flint *et al.* 2008. Nat Rev Microbiol. 6(2):121-131

Bovine Rumen
- ~100 L volume
- Bacteria, Archaea, Protozoa, Viruses, Fungi
- ~1 x 10^{10} bacteria / mL
- Microbes provide genes that cow genome does not encode



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:

- It is well known that biotechnology has a continuous demand for **novel genes** and enzymes and compounds (Christel *et al.*, of the 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.

Table 4. The KEGG pathways with significantly different numbers of hits between the two age groups.

KO#	KO pathway (class)	14d[a]	42d[a]	*P*-value[b]
00230	Purine metabolism (M)	0.0185 ± 1.13E-04	0.0191 ± 3.89E-05	0.0239
00400	Phe, Tyr and Try biosynthesis (M)	0.0056 ± 1.16E-04	0.0062 ± 7.58E-05	0.0429
03070	Bacterial secretion system (E)	0.0053 ± 3.37E-07	0.0054 ± 1.46E-06	0.0000
00300	Lysine biosynthesis (M)	0.0047 ± 1.85E-05	0.0053 ± 3.59E-05	0.0034
00363	Bisphenol degradation (M)	0.0041 ± 4.59E-06	0.0044 ± 7.24E-05	0.0328
00670	One carbon pool by folate (M)	0.0026 ± 3.06E-05	0.0029 ± 1.13E-05	0.0057
00780	Biotin metabolism (M)	0.0021 ± 6.48E-05	0.0025 ± 5.58E-05	0.0417
00401	Novobiocin biosynthesis (M)	0.0019 ± 4.44E-05	0.0022 ± 2.70E-05	0.0157
03410	Base excision repair (G)	0.0017 ± 1.27E-05	0.0018 ± 2.96E-05	0.0227
00950	Isoquinoline alkaloid biosynthesis (M)	0.0015 ± 1.92E-05	0.0019 ± 5.67E-05	0.0215
00514	O-glycan biosynthesis (M)	0.0001 ± 1.49E-06	0.0002 ± 9.80E-06	0.0192

a. The number of open reading frames (ORF) annotated to a given KO pathway divided by the total number of ORF annotated to all KO pathways (mean ± SEM).
b. *P*-value was calculated using MetaStats (White *et al.*, 2009).
M, Metabolism; E, Environmental Information Processing; G, Genetic Information Processing.

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

Rumen microbes	Stain and coverage	Institution	websites
Fibrobacter succinogenes	Strain S 85 closed	TIGR	www.tigr.org/tdb/rumenomics
Ruminococcus albus	Strain 8 closed	TIGR	www.tigr.org/tdb/rumenomics
Prevotella ruminicola	Strain 23 closed	TIGR	www.tigr.org/tdb/rumenomics
Prevotella bryanti	Strain B, 4, Draft (8x)	TIGR	www.tigr.org/tdb/rumenomics
Ruminococcus flavefaciens	Strain FD -1 Draft (2x)	University of Illinois	www.biotech.uiuc.edu

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

Authors	Enzymes/microbial studies	Source
Ferrer *et al.*, 2005	Novel enzyme RA.04 belonging to the alpha-amylase family	Bovine rumen
Beloqui *et al.*, 2006	Novel gene Rl-5 encoding polyphenol oxidase	Bovine rumen
Matthew *et al.*, 2007	Novel methanogens	Cattle, Sheep rumen
Lammie *et al.*, 2007	Rumen microbes	Bovine rumen (SSH)
Palackal *et al.*, 2007	multifunctional glycosyl hydrolase	Bovine rumen

Identification of novel enzymes/microbes from rumen:

- 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).

Science AAAS

Metagenomic Discovery of Biomass-Degrading Genes and Genomes from Cow Rumen
Matthias Hess, *et al.*
Science **331**, 463 (2011);
DOI: 10.1126/science.1200387

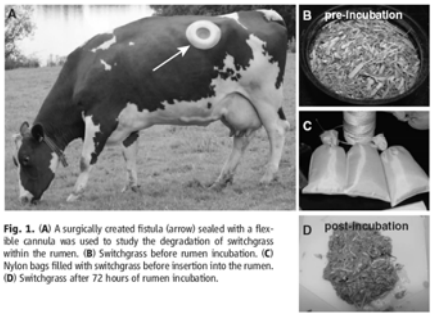


Fig. 1. (**A**) A surgically created fistula (arrow) sealed with a flexible cannula was used to study the degradation of switchgrass within the rumen. (**B**) Switchgrass before rumen incubation. (**C**) Nylon bags filled with switchgrass before insertion into the rumen. (**D**) Switchgrass after 72 hours of rumen incubation.

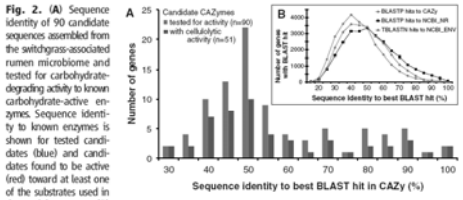


Fig. 2. (**A**) Sequence identity of 90 candidate sequences assembled from the switchgrass-associated rumen microbiome and tested for carbohydrate-degrading activity to known carbohydrate-active enzymes. Sequence identity to known enzymes is shown for tested candidates (blue) and candidates found to be active (red) toward at least one of the substrates used in the activity assays. (**B**) Similarity distribution of CAZyme candidates (n = 27,755) containing a catalytic domain (CD) associated with carbohydrolytic activity or a carbohydrate-binding module (CBM). Sequences were compared to the CAZy (blue, 25,947 hits), NCBI-nr (black, 26,679 hits), and NCBI-env (green, 26,030 hits) databases (best BLAST hit, E-value ≤ 1e-5); 482 genes contained both a CD and CBM, whereas 23,804 and 3469 genes contained only a CD or CBM, respectively.

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

Authors	Enzymes/microbial studies	Source
Ferrer *et al.*, 2005	Novel enzyme RA.04 belonging to the alpha-amylase family	Bovine rumen
Beloqui *et al.*, 2006	Novel gene Rl-5 encoding polyphenol oxidase	Bovine rumen
Matthew *et al.*, 2007	Novel methanogens	Cattle, Sheep rumen
Lammie *et al.*, 2007	Rumen microbes	Bovine rumen (SSH)
Palackal *et al.*, 2007	multifunctional glycosyl hydrolase	Bovine rumen

Identification of uncultured methanogens:

- 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).

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.

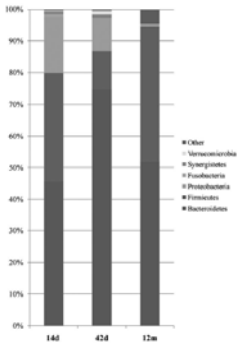


Fig. 1. The relative distribution of the six most abundant phyla in the bovine rumen microbiota of pre-ruminant calves fed milk replacer ($n = 3$). Percentages denote numbers of 16S sequences assigned to a given phylum divided by the total number of 16S sequences that were positively assigned to any phylum. The microbial composition of the rumen microbiota of 12-month-old (12m) bull calves ($n = 4$) fed a hay diet was used as a reference.

Environmental Microbiology (2011)
Metagenomics and the rumen microbiota

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.

Conclusion:

- 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.

What is metagenomics:

- 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 (Kowalchuk *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 ecosystem of human and animals (Lu *et al.*, 2007; Shanks *et al.*, 2006).

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.

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 anti-nutritional Plant Secondary Metabolites (PSMs) such as tannin-polyphenols.