Journal of Applied Microbiology

Journal of Applied Microbiology ISSN 1364-5072

ORIGINAL ARTICLE

Milk microbiome signatures of subclinical mastitis-affected cattle analysed by shotgun sequencing

V.D. Bhatt[1], V.B. Ahir[2], P.G. Koringa[2], S.J. Jakhesara[2], D.N. Rank[3], D.S. Nauriyal[4], A.P. Kunjadia[1] and C.G. Joshi[2]

1 Ashok and Rita Patel Institute of Integrated Study and Research in Biotechnology and Allied Sciences, Anand, Gujarat, India
2 Department of Animal Biotechnology, College of Veterinary Science and A. H., Anand Agricultural University, Anand, Gujarat, India
3 Department of Animal Genetics and Breeding, College of Veterinary Science and A. H., Anand Agricultural University, Anand, Gujarat, India
4 Department of Veterinary Medicine, College of Veterinary Science and A. H., Anand Agricultural University, Anand, Gujarat, India

Keywords
MG-RAST, next-generation sequencing, subclinical mastitis.

Correspondence
Anju P. Kunjadia, Ashok and Rita Patel Institute of Integrated Study and Research in Biotechnology and Allied Sciences, ADIT campus, New Vallabh Vidyanagar 388 120, Anand, Gujarat, India. E-mails: anjuaribas@gmail.com; bhatt_vbhv@yahoo.co.in

2011/1592: received 20 September 2011, revised 8 December 2011 and accepted 28 December 2011

doi:10.1111/j.1365-2672.2012.05244.x

Abstract

Aims: Metagenomic analysis of milk samples collected from Kankrej, Gir (*Bos indicus*) and crossbred (*Bos taurus* × *B. indicus*) cattle harbouring subclinical mastitis was carried out by next-generation sequencing 454 GS-FLX technology to elucidate the microbial community structure of cattle milk.

Methods and Results: Milk samples from Kankrej, Gir and crossbred cattle were subjected to metagenomic profiling by pyrosequencing. The Metagenomic analysis produced 63·07, 11·09 and 7·87 million base pairs (Mb) of sequence data, assembled in 264 798, 56 114 and 36 762 sequences with an average read length of 238, 197 and 214 nucleotides in Kankrej, Gir and crossbred cattle, respectively. Phylogenetic and metabolic profiles by the web-based tool MG-RAST revealed that the members of Enterobacteriales were predominant in mastitic milk followed by Pseudomonadales, Bacillales and Lactobacillales. Around 56 different species with varying abundance were detected in the subclinically infected milk. *Escherichia coli* was found to be the most predominant species in Kankrej and Gir cattle followed by *Pseudomonas aeruginosa*, *Pseudomonas mendocina*, *Shigella flexneri* and *Bacillus cereus*. In crossbred cattle, *Staphylococcus aureus* followed by *Klebsiella pneumoniae*, *Staphylococcus epidermidis* and *E. coli* were detected in descending order. Metabolic profiling indicated fluoroquinolones, methicillin, copper, cobalt–zinc–cadmium as the groups of antibiotics and toxic compounds to which the organisms showed resistance. Sequences indicating potential of organisms exhibiting multidrug resistance against antibiotics and resistance to toxic compounds were also present. Interestingly, presence of bacteriophages against *Staph. aureus*, *E. coli*, *Enterobacter* and *Yersinia* species was also observed.

Conclusions: The analysis identified potential infectious organisms in mastitis, resistance of organisms to antibiotics and chemical compounds and the natural resistance potential of dairy cows.

Significance and Impact of the Study: The findings of this study may help in formulating strategies for the prevention and treatment of mastitis in dairy animals and consequently in reducing economic losses incurred because of it.

Introduction

Mastitis is an inflammatory reaction of the mammary gland generally caused by micro-organisms. The infective agents implicated to cause bovine mastitis include bacteria, fungi and algae. So far, the role of virus, if any, has not been elucidated in causation of mastitis. Flakes or clots in the milk and swelling of infected quarters are

common signs of clinical mastitis; on the other hand, subclinical mastitis, in absence of any clinical sign, usually remains unnoticed. As a result of subclinical mastitis-affected animals produce reduced amounts of milk compared to their true production potential. Thus, subclinical mastitis in dairy cattle can result in reduced milk production with economic consequences for the farmer. Mastitis is regarded as the costliest disease of dairy industry worldwide; it is associated with economic losses of $35 billion annually (Annapoorani *et al.* 2007). Controlling subclinical mastitis can reduce the losses in milk production substantially.

Every dairy herd has cows with subclinical mastitis, and the prevalence of infected cows varies from 15 to 75%, whereas the involvement of quarters differs between 5 and 40% (Cynthia 2005). A variety of pathogens can establish chronic infection that only occasionally manifests clinical signs of mastitis. The primary focus of most subclinical mastitis programmes is to reduce the prevalence of contagious pathogens viz. *Streptococcus agalactiae*, *Staphylococcus aureus* and other Gram-positive cocci, most notably *Streptococcus dysgalactiae* (which may be contagious or be environmentally acquired) and environmental pathogens including *Streptococcus uberis*, *Enterococcus* and numerous other coagulase-negative staphylococci, including *Staphylococcus hyicus*, *Staphylococcus epidermidis*, *Staphylococcus xylosus* and *Staphylococcus intermedius*. This list is not complete and organisms that cannot be cultured by standard culture conditions may escape notification. Metagenomic study provides an opportunity to record such pathogens that can help in planning effective therapeutic and preventive measures.

The global microbial diversity presents an enormous, largely untapped genetic and biological pool (Cowan 2000). In spite of their obvious importance, very little is known about environmental microbes or their diversity; for example, how many species are present in the environment and what their ecological functions are (Singh *et al.* 2008). Until recently, there were no appropriate techniques available to answer such questions because of the limitations encountered in the culturing of microbes; traditional methods of culturing micro-organisms only detect those organisms that grow under laboratory conditions (Hugenholtz *et al.* 1998; Rondon *et al.* 2000). It is widely accepted that up to 99% of the microbes in the environment cannot be readily cultivated (Kamagata and Tamaki 2005; Sekiguchi 2006). To overcome these difficulties and limitations associated with cultivation techniques, different DNA-based molecular methods have been developed for characterizing microbial species and assemblages, and these have significantly influenced our understanding of microbial diversity and ecology (Delong 2005).

Although much of metagenomics focuses on bacteria (especially the 16S rRNA gene), the field is expanding rapidly to encompass the entire spectrum of organisms in an environmental sample that includes bacteria, archaea, viruses, small eukaryotes, plasmids and short RNAs (William 2010). 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 different microbes within the community and the genetic information they contain about microbial niches (Streit and Schmitz 2004). Thus, metagenomics is a rapidly growing field of research that aims to study 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 (Ghazanfar and Azim 2009), and therefore is capable of overcoming these difficulties.

The ongoing revolution in metagenomic sequencing technology has led to the production of sequencing machines with dramatically lower costs and higher throughput. Next-generation platforms are helping to open entirely new areas of biological inquiry, including the investigation into ancient genomes, characterization of ecological diversity and identification of unknown aetiologic agents.

The objective of this study was to undertake a comparative evaluation of microbial community and the metabolic profiling among the milk samples from three breeds of cattle harbouring subclinical mastitis. Studying the DNA and metabolic profiling of microbial communities could help to establish more effective prevention and treatment against potential microbial agents.

Materials and methods

Animals and preparation of milk samples

Milk samples of Kankrej and Gir cows were collected from Livestock Research Station (LRS), and samples of crossbred cattle were collected from Livestock Production and Management (LPM) cattle farms in Anand town of Gujarat State, India. Quarter milk samples were collected aseptically in sterile wide mouth glass stopper bottles on three consecutive days. The udder was washed thoroughly with potassium permanganate solution (1 : 1000), and the teats were wiped with 70% ethyl alcohol prior to sampling.

Quarter milk samples were subjected to cell counts using an electronic somatic cell counter (Foss, Hillerod, Denmark) and to bacteriological culture examination. On the basis of these results, particular quarters were identified

as infected with subclinical mastitis (International Dairy Federation 1987). Ten cows each of crossbred (*Bos taurus* × *Bos indicus*), Kankrej and Gir (*B. indicus*) breeds with subclinically affected quarters were included as experimental material.

Microbiological cultures

All milk samples were evaluated by the 1 μl calibrated loop method for microbiological growth in (i) sheep blood agar medium, (ii) McConkey's medium and (iii) Sabouraud dextrose agar medium. Staining and cellular morphological features of organisms were ascertained by microscopic examination of Gram-stained smears. A standard laboratory protocol of inspecting cultures daily for the growth of potential pathogens was followed.

Total DNA extraction

Thirty millilitres of milk sample from each selected quarters was diluted with 300 ml sterile normal saline solution (pH 7·2). The diluted samples were filtered using 3 micron nitrocellulose filters to remove somatic cells and centrifuged at 16 000 ***g*** for 15 min in 50 ml tubes. The pellet was washed three times with sterile normal saline solution. DNA was extracted using the protocol described by Cremonesi *et al.* (2006). The extracted DNA was resuspended in 1× TE buffer (100 mmol l^{-1} Tris, 10 mmol l^{-1} EDTA, pH 8·0). RNA was removed from the extracts by RNaseA (Qiagen, thilden, Germany) treatment. RNase was removed by phenol: chloroform: isoamyl alcohol (25 : 24 : 1) extraction. The DNA was then precipitated, washed and resuspended in 150 μl TE buffer.

454 sequencing

Quantity and quality of each DNA sample were checked by OD260/280 ratio using an Infinite 200 PRO Nanoquant Spectrophotometer (Tecan, Mannedorf, Switzerland) as well as by visual inspection on 0·8% agarose gel. The DNA samples were pooled breed-wise in equimolar concentration, and 500 ng of total DNA was used for library preparation. The quality and the average size of the library molecules were analysed on the Agilent 2100 Bioanalyzer using the DNA High sensitivity Chip kit (Agilent Technologies, Santa Clara, CA, USA). The breed-wise three libraries were subjected to emulsion PCR. Clonally amplified beads were enriched and subjected to 454 Genome Sequencer FLX (Roche, Branford, CT, USA) run following the manufacturer's protocol.

The raw data were processed by the GS Run Browser to generate assembled reads. The domain distribution data were calculated for the metagenomes of Kankrej, Gir and crossbred cattle and these were compared to SEED (The cooperative effort on the development of the comparative genomics environment: http://www.theseed.org) using a maximum *e*-value of 0·001, a minimum identity of 85% and a minimum alignment length of 20 bp.

Results

Microbiological cultures

The presumptive bacterial isolates from each raw milk sample were further subjected to identification and confirmation. Individual colony characteristics and haemolytic patterns on blood agar were noted. The bacteria isolated were identified on the basis of their cultural, morphological and biochemical characteristics. The results revealed that *Escherichia coli* was present in all the milk samples from Kankrej and Gir cattle. *Staphylococcus aureus* was recovered as the major contagious pathogen in milk from crossbred cattle (Table 1).

Pyrosequencing

The total DNA from milk samples (after removing somatic cells) of Kankrej, Gir and crossbred cattle with subclinical mastitis yielded 63·07, 11·09 and 7·87 million base pairs (Mb) assembled to 264 798, 56 115 and 36 762 sequence reads with an average read length of 238, 197 and 214 nucleotides, respectively.

Table 1 Abundance of bacterial species determined by classical bacteriology, of bacterial genera or species and of bacteriophages detected by Metagenomic analysis

Milk samples	Major bacteria by classical bacteriological analysis	Abundance (%)	Major bacteria by Metagenomics	Abundance (%)	Bacteriophages against major bacteria	Abundance (%)
Kankrej milk	*Escherichia coli*	100	*E. coli*	30·97	*Escherichia* phages	0·04
			Enterobacteria spp.	0·73	Enterobacteria phages	0·28
Gir milk	*E. coli*	100	*E. coli*	26·72	*Escherichia* phages	0·09
			Staphylococcus aureus	11·28	*Staphylococcus* phages	0·44
			Enterobacteria spp.	1·15	Enterobacteria phages	0·52
Crossbred milk	*Staph. aureus*	100	*Staph. aureus*	48·08	*Staphylococcus* phages	2·32

Phylogenetic sequence profile

The distribution of the domains, as calculated from the data, identified bacteria as by far the most dominant (Fig. 1). Phylogenetic profile was carried out using the SEED system databases in MG-RAST (Overbeek *et al.* 2005). SEED-based phylogenetic analyses classified only 4·24% (11 240), 1·99% (1122) and 15·55% (5717) sequences in Kankrej, Gir and crossbred cattle, respectively, while the majority of reads remained nonclassified. It was evident from the comparison that bacteria were abundantly present (>90%), while other classes of microbes were found in minor proportion (<10%). Among bacteria, the major phyla observed were proteobacteria and firmicutes; however, their abundance was different among different breeds of cattle (Fig. 2). The SEED subsystem uses annotated functional genes in a dataset to extract the taxonomic strings associated with their nearest match, thereby resulting in phylogenetic sequence profile. Among the classified sequences that were differentiated into different microbial classes, the domain 'bacteria' contributed an overwhelmingly high proportion of 95·97, 90·47 and 94·24% in Kankrej, Gir and crossbred cattle, respectively.

The order-wise distribution of metagenome sequences showed Aeromonadales, Enterobacteriales, Pseudomonadales, Bacillales and Lactobacillales as predominant orders (Fig. 2), whereas family-wise distribution indicated greater presence of Enterobacteriaceae, Staphylococcaceae, Bacillaceae, Streptococcaceae, Enterococcaceae and Pseudomonadaceae (Table 2).

A heatmap dendrogram indicates the similarity/dissimilarity between a group of vectors. Vectors in a dendrogram are usually ordered with respect to their level of similarity; similar vectors are placed next to each other, while more distantly related vectors are placed farther apart. Each square in the heatmap dendrogram represents the abundance level of a single category in a single sample. The dendrogram shown in Fig. 3 indicates the similarity/dissimilarity among Kankrej, Gir and crossbred samples on *X*-axis, with the categories, that is order on the *Y*-axis. The dendrogram revealed Kankrej and Gir samples as similar and joined them to form the first cluster followed by the crossbred cattle as an out-group. Predominance of Aeromonadales, Enterobacteriales, Pseudomonadales, Bacillales and Lactobacillales was conspicuous by different shades of representative squares (Fig. 3).

On most closely related species-level comparison, *E. coli* were found to be predominant in Kankrej (30·97%) and Gir (26·71%), while *Staph. aureus* (48·08%) in crossbred cattle. The metagenome of the milk samples showed that *E. coli* (30·97%), *Pseudomonas aeruginosa* (10·17%), *Pseudomonas mendocina* (7·73%), *Shigella flexneri* (6·14%), *Shigella sonnei* (5·64%), *Shigella boydii* (4·50%), *Bacillus cereus* (5·69%) and *Salmonella enterica* (4·70%) were the major species in Kankrej, whereas *E. coli* (26·71%), *Ps. aeruginosa* (17·72%), *Staph. aureus* (11·28%), *Sh. flexneri* (5·11%), *Sh. sonnei* (4·67%), *Sh. boydii* (3·70%) and *Salm. enterica* (3·17%) had higher involvement in Gir. In crossbred cattle, the major most closely related species found were *Staph. aureus* (48·08%), *Klebsiella pneumoniae* (12·65%), *Staph. epidermidis* (8·89%), *Salm. enterica* (5·94%) and *E. coli* (5·77%) (Table 3). The PathoSystems Resource Integration Center (PATRIC) analysis showed that 5·72% (12 826), 2·46% (1231) and 13·10% (4896) reads aligned with the database of pathogenic micro-organisms in Kankrej, Gir and crossbred cattle, respectively.

Comparative metagenomics of the microbiota in the milk samples of Kankrej, Gir and crossbred cattle showed that 12·00% of the detected organisms were common to all three breeds. Organisms common to Kankrej and Gir represented 23·30%, those common to Kankrej and crossbred were 18·22% and those common to crossbred and Gir were 13·30%. A Venn diagram shown in Fig. 4.

Interestingly, several bacteriophages were also found to be present in the milk metagenome samples (Table 1), especially *Enterobacteria* and *E. coli* phages (0·32%) in Kankrej, *Enterobacteria* and *Staphylococcus* phages (1·23%)

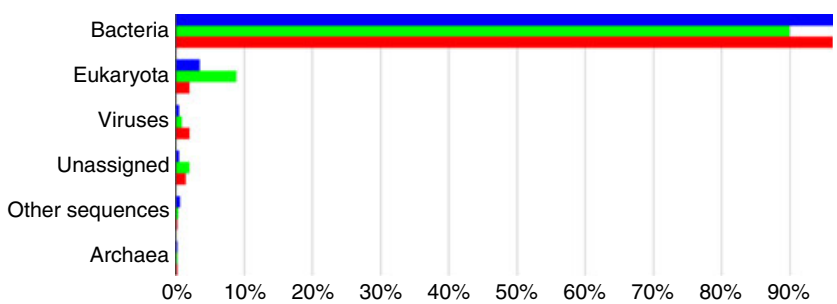

Figure 1 Domain-wise distribution of the sequences identified bacteria as the most common organisms in all three breeds of cattle. (■) Kankrej; (■) Gir and (■) Crossbred.

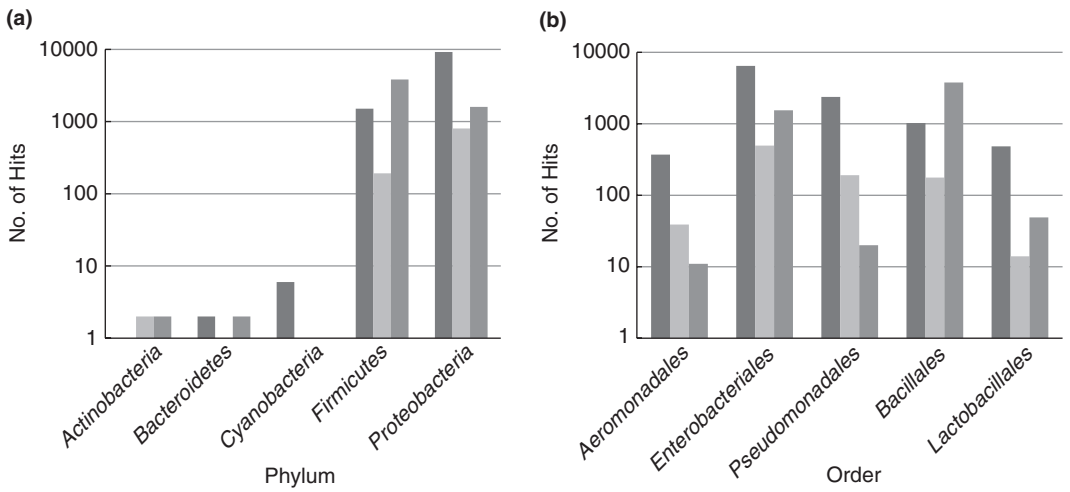

Figure 2 Phylum (a) and order-wise (b) distribution among the bacteria in Kankrej, Gir and crossbred milk metagenome samples. (■) Kankrej; (■) Gir and (■) Crossbred.

in Gir and *Staphylococcus* phages (2·32%) in crossbred cattle.

Metabolic sequence profile

Metabolic profiles were constructed using SEED database that compared the homology of the functional genes against the database and displayed a category of annotated genes with our metagenomic samples. Breed-wise profiling of milk samples by functional classification showed that the 'clustering-based subsystem' was predominant in all samples followed by carbohydrate metabolism, protein metabolism, amino acid and derivatives, RNA metabolism, DNA metabolism, membrane transport

Table 2 Distribution of phylum, order and family among Kankrej, Gir and crossbred metagenome samples

	Kankrej	Gir	Crossbred	Higher classification
Phylum				
Proteobacteria	85·82% (9214)	80·44% (802)	29·39% (1598)	Bacteria
Firmicutes	14·06% (1510)	19·25% (192)	70·49% (3832)	
Other	0·12% (1)	0·31% (2)	0·12% (6)	
Order				
Enterobacteriales	70·11% (6456)	62·9% (502)	96·67% (1543)	Proteobacteria
Pseudomonadales	25·78% (2374)	31·07% (248)	1·25% (20)	
Aeromonadales	4·0% (371)	5·38% (43)	0·0% (0)	
Other	0·11% (13)	0·65% (9)	2·08% (35)	
Bacillales	67·94% (1026)	92·67% (177)	98·72% (3783)	Firmicutes
Lactobacillales	32·05% (484)	7·32% (14)	1·27% (49)	
Other	0·0% (0)	0·0% (0)	0·0% (0)	
Family				
Enterobacteriaceae	100% (6456)	100% (502)	100% (1543)	Enterobacteriales
Bacillaceae	90·54% (929)	5·08% (9)	0·55% (21)	Bacillales
Staphylococcaceae	9·45% (97)	94·91% (168)	99·23% (3754)	
Other	0·0% (0)	0)% (0)	0·22% (8)	
Streptococcaceae	99·17% (480)	28·57% (4)	100% (49)	Lactobacillales
Enterococcaceae	0·41% (2)	71·42% (10)	0% (0)	
Other	0·42% (2)	0·0% (0)	0·0% (0)	
Pseudomonadaceae	100% (2374)	91·90% (228)	95% (19)	Pseudomonadales
Other	0·0% (0)	8·1% (20)	5% (1)	

The data were compared by SEED with a maximum *e*-value of 0·001, a minimum identity of 85% and a minimum alignment length of 20. Numbers of hits are shown in brackets.

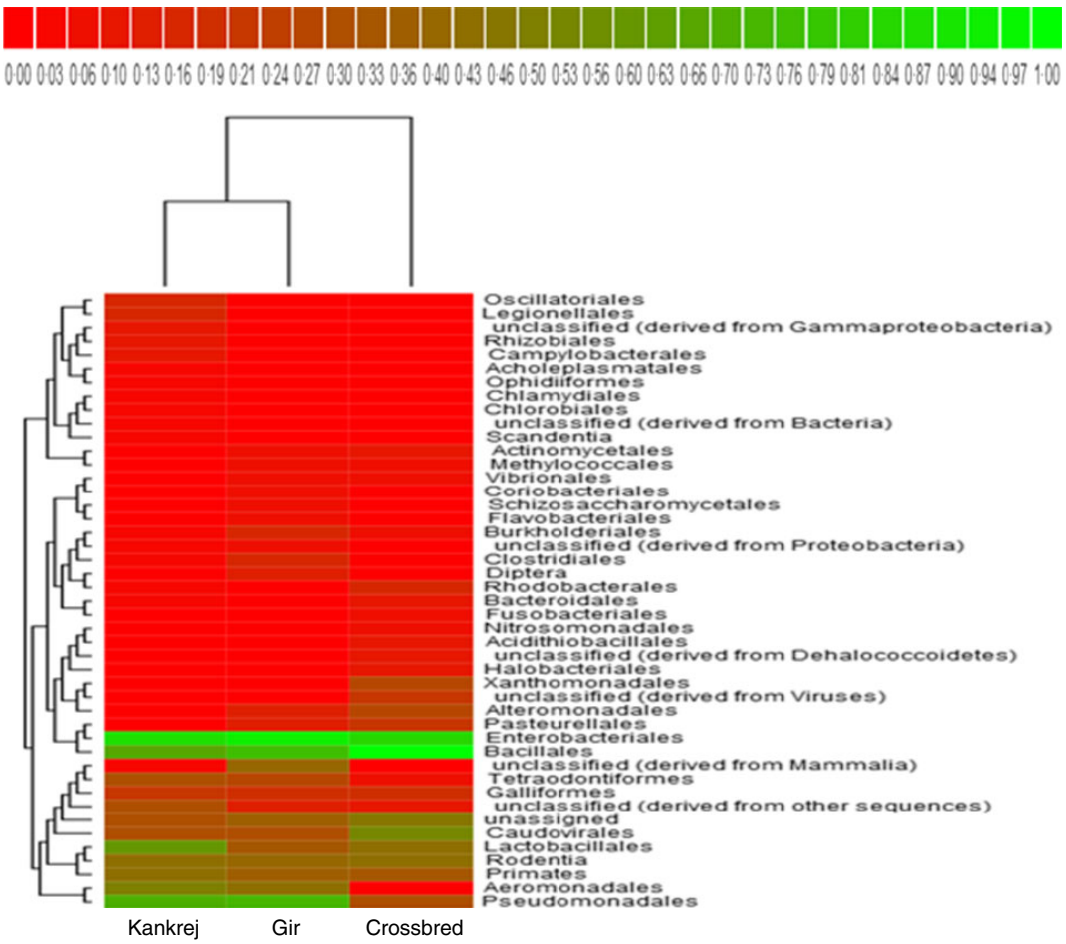

Figure 3 Heatmap dendrogram showing similarities of microbiota composition of milk samples from Kankrej, Gir and crossbred cattle. Kankrej and Gir fall in one cluster with crossbred as out-group. Shades represent effective raw scores from 0 to 1 of all identified sequences.

and virulence, disease and defence-associated genes (Table 4).

Virulence, disease and defence-associated sequences identified by metabolic profiling using MG-RAST tool constituted 2·58% in Kankrej, 3·74% in Gir and 5·54% in crossbred cattle. In this subsystem, resistance to antibiotics and toxic compounds–associated sequences was predominant among all the metagenome samples (Fig. 5). Several sequences belonging to genes that impart resistance against antibiotics and toxic compounds to the organisms were also identified, e.g. resistance to fluoroquinolones, copper and cobalt–zinc–cadmium. Multidrug resistance and methicillin resistance were also represented. Sequences belonging to genes that impart resistance to mercury, arsenic, erythromycin and fosfomycin were also detected in trace amounts (Fig. 6).

Metabolic profiling also revealed the presence of genes associated with phages, prophages, transposable elements and plasmids as 1·80, 2·49 and 4·29% in Kankrej, Gir and crossbred cattle samples, respectively. In this subsystem, Streptococcal phages were the most abundant among all samples. Sequences of the proteins, associated with phage integration and excision, replication and packaging machinery, were observed in all the samples.

Discussion

Recent studies based on genome sequencing data have begun to investigate the true diversity of the microbial world. Microbiota from diverse clinical sources including the psoriatic skin biopsies (Fahlen *et al.* 2011), tissue bacteria in diabetic humans (Amar *et al.* 2011) or the apical root canal (Siqueira *et al.* 2011) has been reported. However, so far, there are no reports on subclinical mastitis cattle milk microbial flora to which we can compare our data.

The present study was aimed to provide an in-depth analysis of microbial succession in subclinical mastitis-

Table 3 Phylogenetic profile of the metagenome sequences based on SEED subsystem and their relative abundance in different milk samples from Kankrej, Gir and crossbred cattle

	Kankrej		Gir		Crossbred	
Organisms	Reads	%	Reads	%	Reads	%
Escherichia coli	3038	30·97	303	26·72	246	5·78
Pseudomonas aeruginosa	998	10·17	201	17·72	0	0
Pseudomonas mendocina	759	7·74	0	0	3	0·07
Shigella flexneri	603	6·15	58	5·11	56	1·31
Bacillus cereus	559	5·7	7	0·62	13	0·31
Shigella sonnei	554	5·65	53	4·67	43	1·01
Salmonella enterica	462	4·71	36	3·17	253	5·94
Shigella boydii	442	4·51	42	3·7	36	0·85
Shigella dysenteriae	233	2·38	21	1·85	0	0
Streptococcus pyogenes	192	1·96	0	0	14	0·33
Bacillus thuringiensis	191	1·95	0	0	1	0·02
Citrobacter koseri	175	1·78	23	2·03	107	2·51
Aeromonas salmonicida	169	1·72	23	2·03	0	0
Aeromonas hydrophila	164	1·67	19	1·68	0	0
Bacillus anthracis	148	1·51	0	0	6	0·14
Streptococcus suis	127	1·29	0	0	0	0
Klebsiella pneumoniae	107	1·09	0	0	539	12·66
Salmonella bongori	85	0·87	8	0·71	36	0·85
Staphylococcus aureus	82	0·84	128	11·29	2048	48·09
Enterobacter sp. 638	72	0·73	13	1·15	108	2·54
Streptococcus agalactiae	58	0·59	1	0·09	2	0·05
Staphylococcus epidermidis	12	0·12	18	1·59	379	8·9
Enterococcus faecalis	1	0·01	6	0·53	0	0
Acinetobacter baumannii	0	0	11	0·97	0	0
Pectobacterium atrosepticum	0	0	8	0·71	0	0
Pseudomonas putida	0	0	15	1·32	4	0·09
Streptococcus uberis	0	0	0	0	30	0·7
Unassigned	31	0·32	20	1·76	0	0
Enterobacteria phage P2	12	0·12	2	0·18	0	0
Enterobacteria phage P4	5	0·05	0	0	0	0
Yersinia phage L-413C	5	0·05	1	0·09	0	0
Enterobacteria phage 933W sensu lato	2	0·02	1	0·09	0	0
Enterobacteria phage lambda	2	0·02	0	0	0	0
Enterobacteria phage PsP3	2	0·02	0	0	0	0
Escherichia Stx1-converting bacteriophage	2	0·02	1	0·09	0	0
Stx2 converting phage II	2	0·02	1	0·09	0	0
Uncultured bacterium	1	0·01	0	0	0	0
Enterobacteria phage BP-4795	0	0	1	0·09	0	0
Enterobacteria phage phiP27	0	0	1	0·09	0	0
Enterobacteria phage SfV	0	0	1	0·09	0	0
Staphylococcus phage 11	0	0	1	0·09	3	0·07
Staphylococcus phage 42E	0	0	1	0·09	9	0·21
Staphylococcus phage 53 sensu lato	0	0	3	0·26	44	1·03
Staphylococcus phage 2638A	0	0	0	0	1	0·02
Staphylococcus phage 3A	0	0	0	0	2	0·05
Staphylococcus phage 55	0	0	0	0	4	0·09
Staphylococcus phage 77	0	0	0	0	5	0·12
Staphylococcus phage phi 12	0	0	0	0	7	0·16
Staphylococcus phage phi13	0	0	0	0	3	0·07
Staphylococcus phage phiETA	0	0	0	0	5	0·12
Staphylococcus phage phiN315	0	0	0	0	3	0·07
Staphylococcus phage phiSLT	0	0	0	0	5	0·12
Staphylococcus phage PVL	0	0	0	0	1	0·02
Staphylococcus phage ROSA	0	0	0	0	6	0·14
Staphylococcus prophage phiPV83	0	0	0	0	1	0·02
Other	241	2·46	56	4·94	194	4·56

The relative abundance of each most closely related species is calculated by finding its proportional hits among total hits obtained.

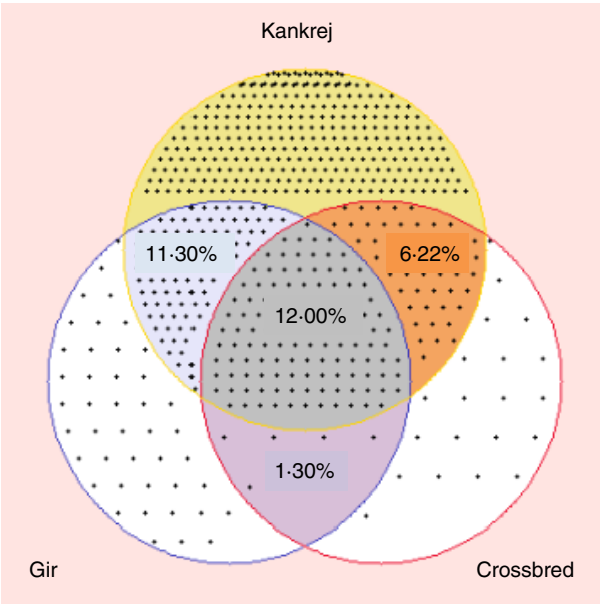

Figure 4 Venn diagram of the organisms shared among the Kankrej, Gir and crossbred milk metagenome.

affected cattle milk utilizing 454 pyrosequencing. The results revealed that Proteobacteria and firmicutes were the main phyla in the milk sampled from three breeds. Proteobacteria are a diverse phylum and include a wide variety of pathogens (Madigan and Martinko 2005). Proteobacteria are Gram-negative and are considered as environmental mastitis pathogens (Hogan *et al.* 1999), while Gram-positive firmicutes are generally considered as contagious mastitis pathogens (Smith and Hogan 1995).

Our findings showed that Kankrej milk had a microbial composition more or less similar to the Gir milk but slightly different from that in the milk of crossbred cows. In the present study, the milk samples from Kankrej and Gir cattle were collected from the same farm, whereas samples from crossbred cattle were from a different farm. Contagious microbial infections can spread among cows or between milk samples during the milking process by contaminated milking equipment, milker's hands, or clothes or sponges used on multiple animals (Fox and Gay 1993). As the Kankrej and Gir cattle included in the present study belonged to the same farm, they were more likely to be infected with the same composition of organisms. The discrimination in the microbial composition of the two groups may be attributed to the genetic makeup of the individual breeds, which could have resulted in differences in susceptibility to different micro-organisms.

Comparison between classical bacteriological culture examination and metagenomic diversity study measurements indicated that in most subclinical cases, there were a handful primary and opportunistic pathogens that made up the largest components of these communities. Thus, subclinical mastitis is a polymicrobial disease, and predominant populations must be considered clinically as potential opportunistic players in such infections. To ignore a potential role of such micro-organisms will limit the potential efficacy of therapy. It is acknowledged that many of the bacteria detected in milk of subclinical mastitis may not be active participants of the community of such bioburdens but may be transient populations or environmental contaminants.

Escherichia coli was the most abundant bacteria in Kankrej and Gir milk metagenome, while *Staph. aureus* had a greater share in crossbred cattle milk. *Escherichia coli* are known to be opportunistic pathogens. There has been a substantial rise in the incidence of *E. coli* mastitis since 1960 and it is known to be the most common cause of fatal mastitis (Menzies *et al.* 1995). *Escherichia coli* is primarily an extracellular pathogen that is involved in intramammary infection (Hill 1991). It is generally accepted that under field conditions, frequent milkout of *E. coli*-infected cows is helpful in the treatment of coliform mastitis. However, when the effect of frequent milkout was tested on the outcome of experimentally induced *E. coli* mastitis in cows, the treatment was shown to be ineffective (Leininger *et al.* 2003) probably because the bacteria replicate very fast.

The next predominant species in the phylum proteobacteria was *Ps. aeruginosa*. The intramammary infection caused by this organism is initially characterized by the development of acute mastitis that often develops into a state of chronic infection (Power 2003). The pathogenicity of *Ps. aeruginosa* is governed by a variety of virulence factors that regulate adhesion, colonization, the production of pili and flagella (Lyczak *et al.* 2000). *Pseudomonas aeruginosa* confers resistance to phagocytosis, facilitates colonization (Lerouge and Vanderleyden 2002).

Antagonism plays an important role in the establishment of any bacterial species in a given niche. Production of antimicrobial substances is one of the strategies applied. Interestingly, *Ps. aeruginosa* produces antimicrobial substances such as pseudomonic acid that inhibits the growth of *Staph. aureus* (Machan *et al.* 1991) and this might explain the predominance of *Ps. aeruginosa* over *Staph. aureus* in the milk of Kankrej and Gir. When *Ps. aeruginosa* was not found, as in crossbred cattle milk, *Staph. aureus* was detected in abundance. *Staphylococcus aureus* is one of the most important aetiologic agents in mastitis of cows, goats and sheep (Mark *et al.* 2005). Moreover, *Staph. aureus* is known to cause chronic and deep infection in the mammary glands that is extremely difficult to cure (Miles *et al.* 1992).

Klebsiella pneumoniae was the second most predominant species in the crossbred cattle milk metagenome.

Table 4 Metabolic profiles of the metagenome sequences and their relative abundance. The latter is calculated by finding its proportional hits among total hits obtained

	Kankrej		Gir		Crossbred	
Subsystems	Hits	Abundance (%)	Hits	Abundance (%)	Hits	Abundance (%)
Clustering-based subsystems	2059	14·14	156	13·88	842	14·04
Carbohydrate metabolism	1884	12·94	147	13·08	712	11·87
Protein metabolism	1262	8·67	68	6·05	351	5·85
Amino acids and derivatives	1222	8·39	100	8·9	441	7·35
Cofactors, vitamins, prosthetic groups, pigments	797	5·47	67	5·96	326	5·44
RNA metabolism	666	4·57	20	1·78	201	3·35
Cell wall and capsule	642	4·41	51	4·54	298	4·97
DNA metabolism	604	4·15	43	3·83	280	4·67
Membrane transport	543	3·73	60	5·34	290	4·84
Nucleosides and nucleotides	490	3·36	29	2·58	161	2·69
Stress response	460	3·16	31	2·76	127	2·12
Respiration	439	3·01	28	2·49	131	2·18
Virulence, disease and defence	375	2·58	42	3·74	332	5·54
Regulation and cell signalling	322	2·21	29	2·58	168	2·8
Fatty acids, lipids and isoprenoids	295	2·03	35	3·11	151	2·52
Phages, prophages, plasmids and transposable elements	262	1·8	28	2·49	257	4·29
Cell division and cell cycle	219	1·5	10	0·89	90	1·5
Nitrogen metabolism	185	1·27	13	1·16	67	1·12
Motility and chemotaxis	158	1·08	11	0·98	22	0·37
Metabolism of aromatic compounds	148	1·02	22	1·96	53	0·88
Sulfur metabolism	147	1·01	13	1·16	50	0·83
Iron acquisition and metabolism	113	0·78	21	1·87	123	2·05
Phosphorus metabolism	112	0·77	13	1·16	24	0·4
Secondary metabolism	72	0·49	3	0·27	23	0·38
Potassium metabolism	62	0·43	5	0·44	34	0·57
Dormancy and sporulation	53	0·36	5	0·44	11	0·18
Photosynthesis	4	0·03	0	0	0	0
Miscellaneous	968	6·65	74	6·58	431	7·19

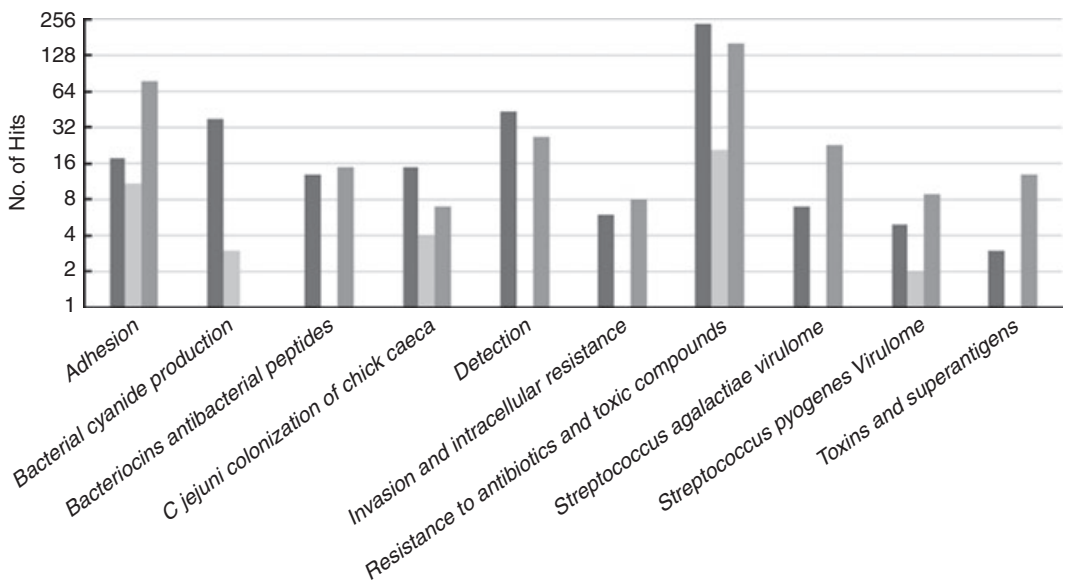

Figure 5 Distribution of 'virulence, disease and defence'-associated sequences in the three metagenome samples. (■) Kankrej; (■) Gir and (■) Crossbred.

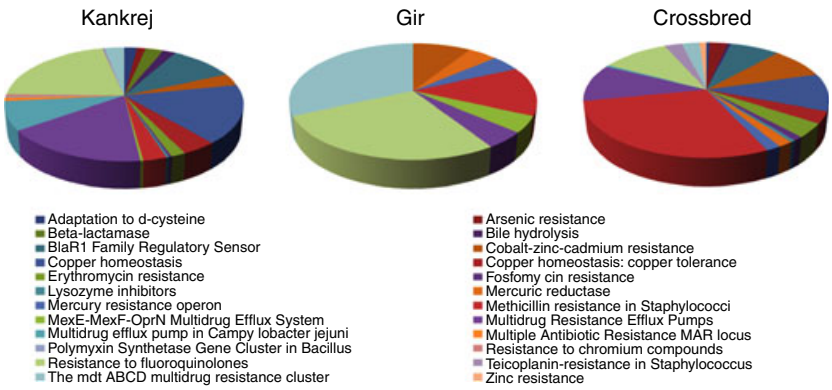

Figure 6 Distribution of 'resistance to antibiotics and toxic compounds'-associated sequences identified resistance to fluoroquinolones, copper, cobalt–zinc–cadmium as well as multidrug resistance and methicillin resistance in the three milk metagenomes.

Like *E. coli*, *Klebsiella* is a Gram-negative environmental pathogen and results in mastitis which can be fatal. In contrast to *E. coli*, *Klebsiella* mastitis can be both clinical and subclinical and responds very poorly to treatment. Cows' faeces are a common source of a large variety of *Klebsiella* strains (Munoz *et al.* 2006; Munoz and Zadoks 2007). We noticed simultaneous presence of *E. coli* and *Klebsiella* in a negligible proportion in Kankrej and Gir cattle.

To the mentioned pathogenic organisms observed in the present study, we can add *Ps. mendocina*, *B. cereus*, *Bacillus thuringiensis*, *Rickettsia prowazekii* and other organisms, which may not commonly be associated with mastitis cases; however, cattle with a compromised immune system are often at risk of developing infection by such organisms.

This is probably the first report describing the presence of bacteriophages in the milk of subclinical mastitic cattle; phages against *Escherichia*, *Enterobacteria*, *Staphylococcus* and *Yersinia* were detected. Interestingly, simultaneous presence of phages along with their hosts in the infected mammary gland was noticed. Our findings suggest that phages may be involved in imparting natural resistance of the cattle against pathogens. Use of *Staphylococcus* phages in the treatment of subclinical *Staph. aureus* mastitis has earlier been reported (Gill *et al.* 2006). In that study involving therapy against established *Staph. aureus* intramammary infections, bacteriological cure rate of 16·7% has been reported. Further, two groups have reported isolation of bacteriophages from the sewage samples and checked their activity against common mastitis pathogens by spot test (Yasunori *et al.* 2008; Aidan *et al.* 2009). The power of 454 pyrosequencing showed detailed pattern of distribution of bacteria and bacteriophages in the mastitic milk.

The widespread and indiscriminate use of antibiotics has played a significant role in the emergence of resistant bacteria. Metabolic profiling has revealed 61·30, 47·73 and 47·52% of sequences in Kankrej, Gir and crossbred cattle samples, respectively, comprised sequences associated with resistance to antibiotics and toxic compounds. These included ribosomally synthesized antibacterial peptides, bacterial cyanide production and tolerance mechanisms, resistance to fluoroquinones, multidrug resistance, methicillin resistance, copper, mercury and zinc.

An earlier study has shown ampicillin, penicillin, streptomycin and oxytetracycline as the most commonly used antibiotics in the treatment of bovine mastitis in the area of survey (Bhatt *et al.* 2011). Indiscriminate and overuse of streptomycin along with penicillin may have resulted in spread of resistance to these two antibiotics. Frequent use of more than one antibiotic may lead to multidrug resistance in the bacteria. Ciprofloxacin is a synthetic chemotherapeutic antibiotic of the fluoroquinolone group (Nelson *et al.* 2007), which is commonly used in the treatment of enteric fever. Frequent use of this drug causes *Salm. enterica* to develop resistance against ciprofloxacin. In the present study, sequences involved in resistance to fluoroquinolone in Kankrej, Gir and crossbred cattle were 21·16, 27·27 and 10·50%, respectively, whereas proportion of *Salm. enterica* were 4·70, 3·17 and 5·94% in Kankrej, Gir and crossbred cattle, respectively. Metabolic profiling thus presented evidence, suggesting that *Salm. enterica* contains genes that might have conferred resistance to the fluoroquinolone antibiotics.

To summarize, bacteria were the most common organisms, and proteobacteria and firmicutes as the predominant phyla observed in the milk metagenome from Kankrej, Gir and crossbred cattle. Further, phylogenetic profiling indicated that Kankrej and Gir metagenome were closer to each other than with crossbred cattle. *Escherichia coli*, *Ps. aeruginosa*, *Sh. flexneri*, *Sh. sonnei* and *Sh. boydii* were the major bacterial species in Kankrej and Gir, while *Staph. aureus*, *Kl. pneumoniae* and *Staph. epidermidis* in crossbred cattle. Culture-based diagnostics

remain difficult especially in case of polymicrobial infections because not all organisms causing infection are recovered on culture. It was demonstrated in our study that subclinical mastitis is not merely caused by a single pathogenic species of bacteria, but rather by a blend of several microbes. In the absence of a true profile of causative organisms, it is quite likely that the potential role of such micro-organisms is ignored, which, in turn, might limit the efficacy of any therapeutic regimen that targets the microbial bioburden.

Acknowledgements

This work was supported by a research grant from Rashtriya Krishi Vikas Yojana (National Agricultural Development Program), Government of India, New Delhi. We are thankful to AMUL for providing the facility of electronic somatic cell counter.

References

Aidan, J.S., Ying, K., Miki, K., Keiko, Y., Hidetomo, I. and Yasunori, T. (2009) Isolation from sewage influent and characterization of novel *Staphylococcus aureus* bacteriophages with wide host ranges and potent lytic capabilities. *Appl Environ Microbial* **75**, 4483–4490.

Amar, J., Serino, M., Lange, C., Chabo, C., Iacovoni, J., Mondot, S., Lepage, P., Klopp, C. *et al.* (2011) Involvement of tissue bacteria in the onset of diabetes in humans: evidence for a concept. *Diabetologia* **54**, 3055–3061.

Annapoorani, C., Dante, S. and Douglas, D. (2007) Antimicrobial activity of bovine bactericidal permeability-increasing protein-derived peptides against gram-negative bacteria isolated from the milk of cows with clinical mastitis. *Am J Vet Res* **68**, 1151–1159.

Bhatt, V.D., Patel, M.S., Joshi, C.G. and Kunjadia, A.P. (2011) Identification and antibiogram of microbes associated with bovine mastitis. *Anim Biotechnol* **22**, 163–169.

Cowan, D.A. (2000) Microbial genomes-the untapped resource. *Trends Biotechnol* **18**, 14–16.

Cremonesi, P., Castiglion, B., Malferrari, G., Biunno, I., Vimercati, C. and Moroni, P. (2006) Technical note: improved method for rapid DNA extraction of mastitis pathogens directly from milk. *J Dairy Sci* **89**, 163–169.

Cynthia, M.K. (2005) *The Merck Veterinary Manual*. Merck and CO., Inc: Whitehouse Station, NJ.

Delong, E.F. (2005) Microbial community genomics in the ocean. *Nat Rev Microbiol* **3**, 459–469.

Fahlen, A., Engstrand, L., Barbara, S., Anne, P. and Lionel, F. (2011) Comparison of bacterial microbiota in skin biopsies from normal and psoriatic skin. *Arch Dermatol Res* **304**, 15–22.

Fox, L.K. and Gay, J.M. (1993) Contagious mastitis. *Vet Clin North Am Food Anim Pract* **9**, 475–487.

Ghazanfar, S. and Azim, A. (2009) Metagenomics and its application in rumen ecosystem: potential biotechnological prospects. *Pak J Nutr* **8**, 1309–1315.

Gill, J.J., Pacan, J.C., Carson, M.E., Leslie, K.E., Griffiths, M.W. and Sabour, P.M. (2006) Efficacy and pharmacokinetics of bacteriophage therapy in treatment of subclinical *Staphylococcus aureus* mastitis in lactating dairy cattle. *Antimicrob Agents Chemother* **50**, 2912–2918.

Hill, A.W. (1991) Somatic cells – friends or foes? In *New Insights into the Pathogenesis of Mastitis* ed. Burvenich, C., Vandeputte-Van Messom, G. and Hill, A.W. *Flemish Vet J*. pp. 217–232.

Hogan, J.S., Gonzalez, R.N., Harmon, R.J., Nickerson, S.C., Oliver, S.P., Pankey, J.W. and Smith, K.L. (1999) *Laboratory Handbook on Bovine Mastitis*. Madison, WI: National Mastitis Council, Inc..

Hugenholtz, P., Goebel, B.M. and Pace, N.R. (1998) Impact of culture independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* **180**, 4765–4774.

International Dairy Federation (1987) Definition and guidelines for diagnosis of bovine mastitis. IDF- Bulletin No. 258.

Kamagata, Y. and Tamaki, H. (2005) Cultivation of uncultured fastidious microbes. *Microbes Environ* **20**, 85–91.

Leininger, D.J., Roberson, J.R., Elvinger, F., Ward, D. and Akers, R.M. (2003) Evaluation of frequent milkout for treatment of cows with experimentally induced *Escherichia coli* mastitis. *J Am Vet Med Assoc* **222**, 63–66.

Lerouge, I. and Vanderleyden, J. (2002) O-antigen structural variation: mechanisms and possible roles in animal/plant–microbe interactions. *FEMS Microbiol Rev* **26**, 17–47.

Lyczak, J.B., Cannon, C.L. and Pier, G.B. (2000) Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes Infect* **2**, 1051–1060.

Machan, Z.A., Pit, T.L., White, W., Watson, D., Taylor, G.W., Cole, P.J. and Wilsons, R. (1991) Interaction between *Pseudomonas aeruginosa* and *Staphylococcus aureus*: description of an antistaphylococcal substance. *J Med Microbiol* **34**, 213–217.

Madigan, M.T. and Martinko, J.M. (2005) *Brock Biology of Microorganisms*, 11th edn. NJ, USA: Prentice Hall.

Mark, T., Tollersrud, T., Kvitle, B., Jørgensen, H.J. and Waage, S. (2005) Comparison of *Staphylococcus aureus* genotypes recovered from cases of bovine, ovine, and caprine mastitis. *J Clin Microbiol* **43**, 3979–3984.

Menzies, F.D., Bryson, D.G., McCallion, T. and Matthews, D.I. (1995) A study of mortality among suckler and dairy cows in Northern Ireland in 1992. *Vet Rec* **137**, 531–536.

Miles, H., Lesser, W. and Sears, P. (1992) The economic implications of bioengineered mastitis control. *J Dairy Sci* **75**, 596–605.

Munoz, M.A. and Zadoks, R.N. (2007) Patterns of fecal shedding of *Klebsiella* by dairy cows. *J Dairy Sci* **90**, 1220–1224.

Munoz, M.A., Ahlstrom, B., Rauch, B.J. and Zadoks, R.N. (2006) Fecal shedding of *Klebsiella pneumoniae* by dairy cows. *J Dairy Sci* **89**, 3425–3430.

Nelson, J.M., Chiller, T.M., Powers, J.H. and Angulo, F.J. (2007) Fluoroquinolone-resistant Campylobacter species and the withdrawal of fluoroquinolones from use in poultry: a public health success story. *Clin Infect Dis* **44**, 977–980.

Overbeck, R., Tadhg, B., Ralph, B., Jomuna, V.C., Han-Yu, C., Matthew, C., Valérie, L., Naryttza, D. *et al.* (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucl Acids Res* **33**, 5691–5702.

Power, E. (2003) Gangrenous mastitis in dairy herds. *Vet Rec* **153**, 791–792.

Rondon, M.R., August, P.R., Bettermann, A.D., Brady, S.F., Grossman, T.H., Liles, M.R. *et al.* (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol* **66**, 2541–2547.

Sekiguchi, Y. (2006) Yet-to-be cultural microorganisms relevant to methane fermentation processes. *Microbes Environ* **21**, 1–15.

Singh, B., Gautam, S.K., Verma, V., Kumar, M. and Singh, B. (2008) Metagenomics in animal gastrointestinal ecosystem: potential biotechnological prospects. *Anaerobe* **14**, 138.

Siqueira, J.F., Flávio, R.F. and Isabela, N.R. (2011) Pyrosequencing analysis of the apical root canal microbiota. *J Endod* **37**, 1499–1503.

Smith, K.L. and Hogan, J.S. (1995) Epidemiology of mastitis. Proc. 3rd Int. Mastitis Seminar, Tel Aviv, Israel **S6**, 3–12.

Streit, W.R. and Schmitz, R.A. (2004) Metagenomics, the key to the uncultured microbes. *Curr Opin Microbiol* **7**, 492–498.

William, J.J. (2010) High-throughput sequencing and metagenomics. *Estuar Coast* **33**, 944–952.

Yasunori, T., Kenji, H., Kohichi, S. and Kazuhiko, M. (2008) Spontaneous deletion of a 209-kilobase-pair fragment from the *Escherichia coli* genome occurs with acquisition of resistance to an assortment of infectious phages. *Appl Environ Microbial* **74**, 4256–4263.