The Journal of Microbiology, October 2008, p. 469-477 Vol. 46, No. 5
DOI 10.1007/s12275-008-0117-z

Turkey Fecal Microbial Community Structure and Functional Gene Diversity Revealed by 16S rRNA Gene and Metagenomic Sequences

Jingrang Lu and Jorge Santo Domingo*

U.S. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, MS-387, Cincinnati, OH, USA

(Received May 10, 2008 / Accepted July 16, 2008)

The primary goal of this study was to better understand the microbial composition and functional genetic diversity associated with turkey fecal communities. To achieve this, 16S rRNA gene and metagenomic clone libraries were sequenced from turkey fecal samples. The analysis of 382 16S rRNA gene sequences showed that the most abundant bacteria were closely related to *Lactobacillales* (47%), *Bacillales* (31%), and *Clostridiales* (11%). *Actinomycetales*, *Enterobacteriales*, and *Bacteroidales* sequences were also identified, but represented a smaller part of the community. The analysis of 379 metagenomic sequences showed that most clones were similar to bacterial protein sequences (58%). Bacteriophage (10%) and avian viruses (3%) sequences were also represented. Of all metagenomic clones potentially encoding for bacterial proteins, most were similar to low G+C Gram-positive bacterial proteins, particularly from *Lactobacillales* (50%), *Bacillales* (11%), and *Clostridiales* (8%). Bioinformatic analyses suggested the presence of genes encoding for membrane proteins, lipoproteins, hydrolases, and functional genes associated with the metabolism of nitrogen and sulfur containing compounds. The results from this study further confirmed the predominance of *Firmicutes* in the avian gut and highlight the value of coupling 16S rRNA gene and metagenomic sequencing data analysis to study the microbial composition of avian fecal microbial communities.

***Keywords*:** metagenomics, 16S rRNA gene, Turkey Feces

Due to the difficulties and limitations associated with cultivation techniques, recent studies examining the composition of fecal microbial communities have focused on the analysis of 16S rRNA gene clone libraries. Most studies have focused on studying mammalian systems. Examples include the fecal microbial communities of ruminants (Edwards *et al.*, 2004), horses (Daly *et al.*, 2001), pigs (Leser *et al.*, 2002; Cotta *et al.*, 2003), and humans (Suau *et al.*, 1999; Eckgurg *et al.*, 2005), among others. Overall, these studies have shown that most mammals harbor a diverse microbial community, in which uncultivated members of the *Bacteroidetes* and *Firmicutes* are numerically dominant. In avian species, most studies have focused on specific bacterial populations or pathogens (Collins *et al.*, 2002; Marois *et al.*, 2002). In fact, only a few studies have characterized in some detail the avian fecal bacterial community using sequencing approaches. Lu *et al.* (2003) studied the microbial composition of poultry litter and found that 82% of the total sequences were identified as Gram-positive bacteria, with 62% of total belonging to low G+C Gram-positive groups. Additional information has been obtained by examining intestinal contents. For example, in chicken cecal contents, 49 to 75% of sequences analyzed are closely related to low G+C Gram-positive bacteria, while *Bacteroides* like sequences represent a smaller fraction (i.e., 5~26%; Zhu *et al.*, 2002; Lu *et al.*, 2003). For other birds information on their gut microbial composition is extremely limited, although recent studies have indicated that in contrast with chicken, *Bacteriodetes* dominate the turkey cecum (Scupham, 2007a).

In spite of their value at identifying prokaryotic taxa and species present in fecal samples, 16S rRNA gene PCR cloning methods can introduce several biases when describing microbial community structure. For example, the heterogeneity between copies of the 16S rDNA in many species (Case *et al.*, 2007), the differences in *rrn* operon copies between species (Crosby and Criddle, 2003; Acinas *et al.*, 2004), the impact of the *in situ* physiological status of a bacterium on the number of genomes (and thus *rrn* operon copies) per cell (Ludwig and Schleifer, 2000; Janssen, 2006), and the inefficient amplification of the targeted 16S rDNA region due to primer mismatches (Suzuki *et al.*, 2000) have important implications when studying diversity and abundance of particular bacterial lineages in natural samples. Additionally, 16S rDNA analyses provide limited information on the genetic potential of different microbes within the community. In contrast, metagenomic sequencing analysis can increase our knowledge on the molecular diversity of functional genes of uncultured microorganisms, and by reducing the biases introduced by 16S rRNA gene PCR cloning techniques, better describe the structure of microbial communities (Rondon *et al.*, 2000; Liles *et al.*, 2003). Indeed, metagenomic approaches performed with fecal samples have generated new information regarding the overall composition and functional attributes of the animal gut microbiota.

* To whom correspondence should be addressed.
(Tel) 1-513-569-7085; (Fax) 1-513-569-7328
(E-mail) santodomingo.jorge@epa.gov

For example, Gill *et al.* (2006) recently showed that many gut bacterial populations are responsible for the metabolism of glycans, amino acids, and xenobiotics, otherwise recalcitrant to humans. Moreover, Breitbart *et al.* (2003) recently documented that the viral community in human feces contains at least 1200 viral genotypes, information that clearly cannot be obtained when using 16S rDNA-based approaches.

The objective of this study was to provide an insight into the composition and genetic diversity of turkey fecal microbial community using both phylogenetic and metagenomic sequencing analysis. The value of combining both approaches to better describe the overall microbial composition and predict the functional roles of particular groups of the turkey fecal microbial community is discussed.

Materials and Methods

Sample collection and DNA extraction

Turkey fecal samples were collected from anonymous farms in Ohio (OH, n=20) and West Virginia (WV, n=11) and kindly provided by Melinda Harris (Ohio-EPA) and Don Stoeckel (OH-USGS). Droppings from healthy animals were aseptically collected immediately after defecation by animals, placed into sterile 50 ml centrifuge tubes, and stored at -80°C until used in DNA extractions. DNA was extracted from the fecal samples using the Fast DNA Kit (Qbiogene, USA) following manufacturer's instructions. Genomic DNA was eluted in 80 μl of water, and the concentration of each fecal DNA extract was measured using a NanoDrop® ND-1000 UV-Vis Spectrophotometer (NanoDrop Technologies, Inc., Germany). The DNA extracts for individual DNA samples were kept at -80°C.

Development of 16S rRNA gene clone library

Sequence analysis of 16S rRNA gene products was used to describe the phylogenetic group affiliation of bacterial populations in turkey fecal extracts. Aliquots from each turkey fecal DNA extracts (i.e., 323 ng per individual sample) were pooled to form a composite DNA sample, which was then used as the PCR template. The 16S rRNA gene was amplified from turkey fecal DNA composite using eubacterial primers 27F; 5'-AGAGTTTGATCMTGGCTCAG-3' and 785R; 5'-ACTACCRGGGTATCTAATCC-3' (Lu *et al.*, 2007). PCR amplifications were performed in a PTC-240 DNA Engine Tetrad 2 Cycler (MJ Research, Inc., USA). Reactions were carried out in a 25 μl volume using the following cycling conditions: 3 min at 95°C, followed by 22 cycles of 30 sec at 95°C, 30 sec at 58°C and 60 sec at 72°C, and a final 10 min primer extension step at 72°C.

Development of turkey fecal metagenomic library

The metagenomic clone library was developed using the methods described by Lu *et al.* (2007). Briefly, the fecal DNA composite was mechanically sheared into approximately 100 to 900 bp fragments and oligonucleotide linkers with degenerate ends were added using polymerase I Klenow fragment (5'-GACACTCTCGAGACATCACCGGTACCNNN NNNNNN-3'; underlined bases represent the linkers used as the priming sites in the single-primer amplification reactions, see below). The modified DNA fragments (i.e., with oligonucleotide linkers) were purified using a QiaQuick PCR Product Clean-up Kit (QIAGEN, USA) and then used in single-primer amplification reactions (50 μl) containing 1× Ex *Taq* PCR buffer, 2.5 mM (each) dATP, dCTP, dGTP, and dTTP, 0.2 μM of PCR primer (5'-GACACTCTCGAGA CATCACCGG-3'), 1% acetamide, 0.625 U Ex *Taq* (TaKaRa Bio Inc., Japan), and 10 μl of processed metagenomic DNA. Incubation temperatures were 94°C (40 sec), 53°C (1 min), and 72°C (30 sec) for 28 cycles, followed by a 72°C (1.5 min) extension step.

DNA sequencing

Both 16S rRNA gene and metagenomic PCR products from five independent reactions were pooled and cloned into pCR4.1 TOPO A Kit (Invitrogen, USA). Individual clones were subcultured in 250 μl of Luria-Bertani broth mixed with 10 μg/ml ampicillin, and aliquots of cultures then used in amplification assays to screen for PCR product inserts. The amplification reactions (25 μl) contained 1× Ex *Taq* PCR buffer, 2.5 mM (each) dATP, dCTP, dGTP, and dTTP, 0.2 μM of M13F; 5'-GTAAAACGACGGCCAG-3' and M13R; 5'-CAGGAAACAGCTATGAC-3' primers, 0.064% Bovine serum albumin (Sigma), 0.625 U Ex *Taq* (TaKaRa, Japan), and 1 μl of template. Incubation temperatures included 94°C (3 min) for the lysis step, followed by 20 cycles of 94°C (30 sec), 52°C (20 sec), and 72°C (40 sec). Prior to sequencing, PCR products were purified using Qiaquick 96 Plate (QIAGEN, USA). Sequencing was carried out at the Cincinnati Children's Hospital Medical Center Genomics Core Facility (Cincinnati, USA) using M13 forward and reverse primers, Big Dye terminator chemistry, and Applied Biosystem PRISM 3730XL DNA Analyzer.

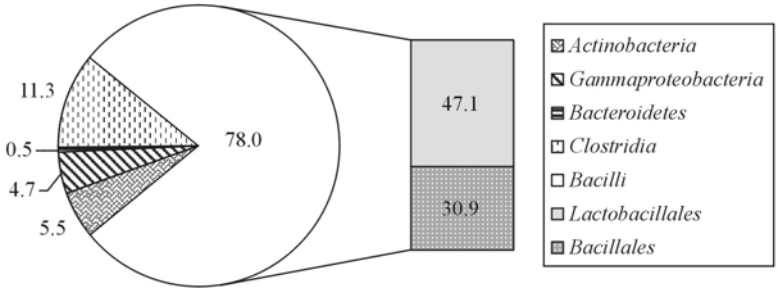

Fig. 1. Bacterial community composition (%) of turkey fecal DNA observed in the 16S rRNA gene clone library. A total of 382 clones were analyzed in this study.

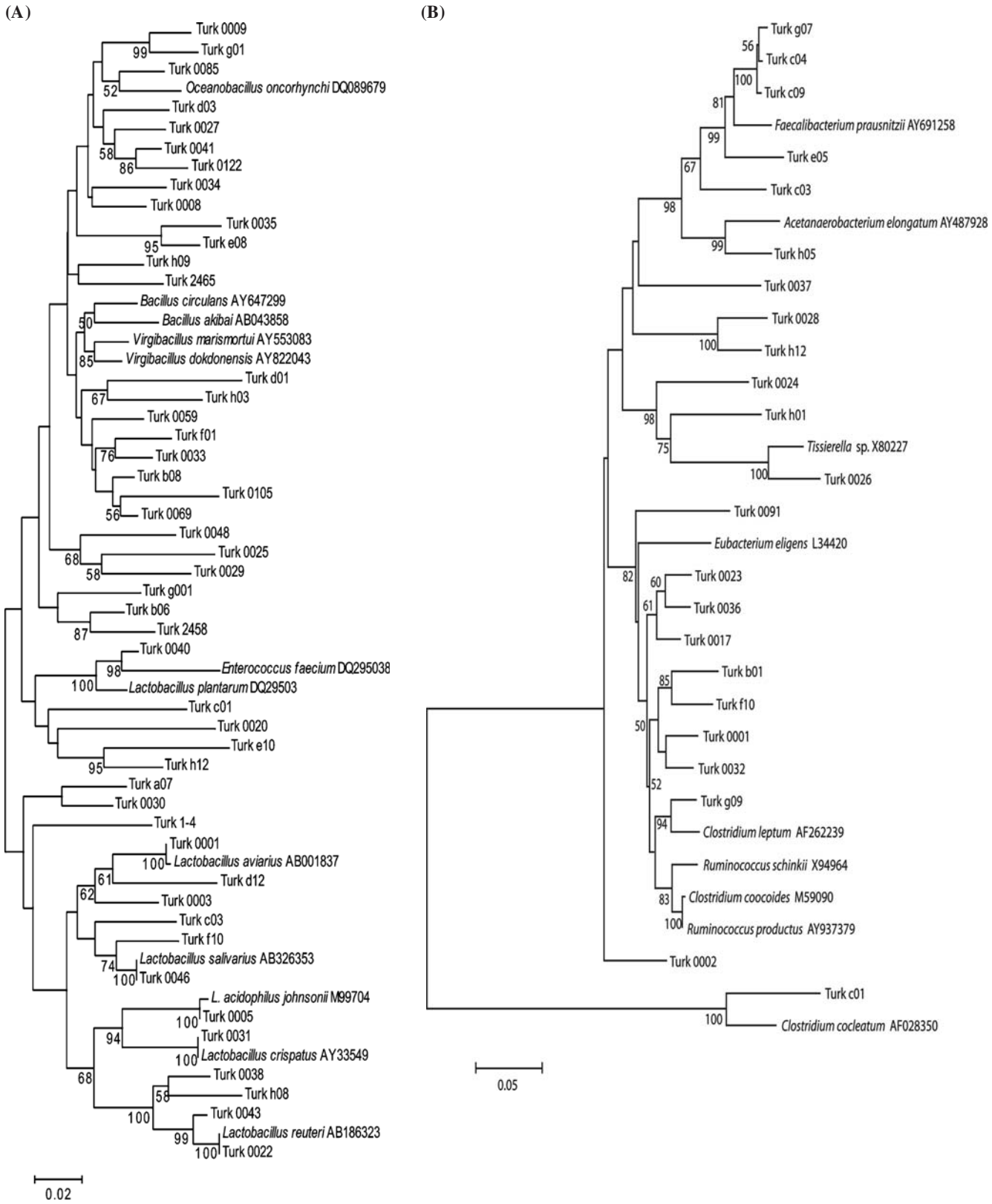

Fig. 2. Phylogenetic trees of 16S rRNA gene turkey fecal DNA clone sequences closely related to *Bacilli* (A, n=290) and *Clostridia* (B, n=43). The phylogenetic trees were developed using Neighbor joining algorithm and MEGA version 3.0 software (Molecular Evolutionary Genetics Analysis software, Arizona State University. Numbers indicate data from a bootstrap analysis (n=1000). Only values above 50% are shown.

Table 1. Representative metagenomic sequences[a]

Sequences ID	Suggested function	Organism (accession no.)	% Identity (protein sequence)
Turkey-membrane1 (EU873890)	membrane protein involved in the export of O-antigen and teichoic acids	*Pediococcus pentosaceu* (ZP 00323775)	33/93 (35%)
Turkey-membrane2 (EU873939)	predicted membrane protein	*Oenococcus oeni* (ZP_00319323)	68/142 (47%)
Turkey-lip1 (EU873944)	putative lipoprotein	*Bacteroides fragilis* (CAH06058)	23/85 (27%)
Turkey-elongation (EU873905)	GTP-binding translation elongation factor LepA	*Lactobacillus plantarum* (CAD64391)	59/78 (75%)
Turkey-chaperone (EU874186)	ATPase with chaperone activity ATP-binding subunit	*Streptococcus suis* (ZP_00333083)	49/72 (68%)
Turkey-hydrolase1 (EU874002)	metal-dependent phosphohydrolase HD subdomain	*Clostridium thermocellum* (P 00504467)	38/79 (48%)
Turkey-hydrolase2 (EU873936)	similar to hydrolase (metallo-β-lactamase superfamily)	*Clostridium beijerincki* (ZP 00910511)	28/77 (36%)
Turkey-esch (EU873916)	putative export protein D for general secretion pathway (GSP)	*Escherichia coli K12* (AAC76350)	94/96 (97%)
Turkey-enter (EU874064)	aminotransferase class I and II	*Enterococcus faecium* (EAN09361)	58/105 (55%)
Turkey-staph (EU874165)	RNA polymerase β-prime chain	*Staphylococcus epidermidis* (NP 763862)	70/88 (79%)
Turkey-bif (EU874118)	putative ribosomal protection protein TetW	*Bifidobacterium* sp. (AAP21776)	93/94 (98%)
Turkey-prop (EU874066)	23S rRNA methyltransferase	*Propionibacterium acnes* (AAL07369)	31/31 (100%)
Turkey-cory (EU873900)	transposase	*Corynebacterium jeikeium* (AAK28908)	32/32 (100%)
Turkey-lact1 (EU873923)	dihydroxyacetone kinase	*Lactobacillus acidophilus* (YP 194289)	72/89 (80%)
Turkey-lact2 (EU874078)	chain B, maltose phosphorylase	*L. brevis* (1H54_B)	76/95 (80%)
Turkey-lact3 (EU874225)	gamma-aminobutyrate permease and related permeases	*L. gasser* (ZP_00341781)	81/88 (92%)
Turkey-lact4 (EU873969)	phosphoglucomutase/phosphomannomutase	*L. johnsonii* (AAS08715)	94/95 (98%)
Turkey-lact5 (EU873967)	thymidylate synthase	*L. johnsonii* (AAS09239)	79/111 (71%)
Turkey-lact6 (EU874203)	fructose-bisphosphate aldolase	*L. plantarum* (NP_784142)	48/55 (87%)
Turkey-lact7 (EU873984)	sucrose PTS transporter	*L. acidophilus* (AAO38866)	94/132 (71%)
Turkey-lact8 (EU874128)	Na^+/H^+ antiporter	*L. plantarum* (NP_7846260	61/99 (61%)
Turkey-lact9 (EU874147)	purine-cytosine transport protein	*L. plantarum* (NP_786604)	31/45 (68%)
Turk-lact10 (EU873920)	glucose-6-phosphate isomerase	*L. sakei* (CAI55481)	80/99 (80%)
Turk-lact11 (EU874076)	diphosphomevalonate decarboxylase	*L. sakei* (YP_395518)	86/131 (65%)
Turk-nitr1 (EU874030)	tyrosine recombinase XerD	*Nitrosococcus oceani* (YP_344243)	25/56 (44%)
Turk-nitr2 (EU874207)	integrase catalytic core	*Nitrosomonas europaea*	63/111 (56%)
Turk-par1 (EU874132)	hypothetical protein PdenDRAFT_4311 P	*Paracoccus denitrificans* (EAN67704)	24/55 (43%)
Turk-arth1 (EU874222)	ribosomal protein S10 bacterial form	*Arthrobacter* sp. (ZP_00412367)	58/62 (93%)
Turk-therm1 (EU873964)	phosphotransferase system HPr-related proteins MB4	*Thermoanaerobacter tengcongensis* (AAM25012)	33/80 (41%)
Turk-syntr1 (EU874075)	510-methylenetetrahydrofolate reductase (FADH)	*Syntrophomonas wolfei* (ZP 00662233)	64/109 (58%)
Turk-desul1 (EU873917)	K^+ transporter Trk	*Desulfuromonas acetoxidans* (ZP 00552576)	41/76 (53%)
Turk-desul2 (EU873935)	hypothetical protein DhafDRAFT_2511	*Desulfitobacterium hafniense* (ZP 00558345)	47/83 (56%)
Turk-noc1 (EU874143)	catalytic LigB subunit of aromatic ring-opening dioxygenase	*Nocardioides* sp. (ZP_00659872)	47/76 (61%)
Turk-brev1 (EU874185)	acyl-CoA dehydrogenases	*Brevibacterium linens* (ZP_0037808)	44/66 (66%)
Turk-alk1 (EU874200)	dihydropteroate synthase DHPS	*Alkaliphilus metalliredigenes* (ZP 00799023)	48/93 (51%)
Turk-anaer1 (EU873918)	hypothetical protein AdehDRAFT_0508	*Anaeromyxobacter dehalogenans*	49/76 (64%)
Turk-phag1 (EU874024)	capsid protein major	*Lactococcus lactis* phage (NP 663677)	76/119 (63%)
Turk-phag2 (EU873907)	terminase large subunit	*Listonella pelagia* phage (YP 224236)	65/117 (55%)
Turk-phag3 (EU874111)	tail tape measure protein TP901 core region	*Alkaliphilus metalliredigenes* phage (ZP 00799480)	86/210 (40%)
Turk-virus1 9 (EU874013)	Rep protein	Avian adeno-associated virus (AAT48612)	111/131 (84%)
Turk-virus2 (EU873897)	pTP	Fowl adenovirus 1 (AAR97561)	99/118 (83%)
Turk-virus3 (EU873993)	capsid protein	Avian adeno-associated virus strain goose parvovirus (YP 077183)	93/101 (92%)

[a] e score $>10^{-10}$

Data analysis

Sequence editing and alignment were completed using Sequencher (Gene Codes Corporation, USA). The 16S rRNA gene sequences were screened for chimeras with the Check Chimera program of the Ribosomal Database Project (Cole *et al*., 2007) and by manual alignments of secondary structure. The Bellerophon program was also used to check for chimeras (Huber *et al.*, 2004). As a final check for chimeras, each sequence was split into equal size 5' and 3' fragments, which were analyzed separately by BLAST searching of GenBank. Sequences for which either the 5' or 3' fragment had significantly different closest relatives were considered probable chimeras and were removed from the data set (Schmeisser *et al.*, 2003). For 16S rRNA gene sequences, similarity searches of DNA sequences were performed using the non-redundant (NR) GenBank database and BLASTn (NR) (Althschul *et al.*, 1997). The putative protein transcript of each metagenome sequence was annotated based on the biochemical function of similar gene sequences using BLASTX with the non-redundant (NR) GenBank database. BLASTX sequence matches with E value of $\leq 10^{-3}$ and sequence identities of ≥30% were considered to be similar protein sequences (Pearl *et al.*, 2000; Breitbart *et al.*, 2003). When several matches fit these criteria, sequences from intestinal bacterial with presumed known functions were selected as the best fit for protein classification.

To organize sequences into functional gene categories, the DNA sequences were grouped according to the Clusters of Orthologous Groups (COG) of proteins database (http://www.ncbi.nlm.nih.gov/COG; Tatusov *et al.*, 1997). Cloned sequences were assigned to bacterial class annotations based on the top BLASTX hit (lowest e-value score) in the GenBank NR database. Sequences obtained in this study were deposited in GenBank under accession numbers EU873553-EU874240.

Results

From the total of 382 16S rRNA gene sequences analyzed there were 91 operational taxonomic units (OTUs, defined as having 97% sequence identity), representing 76% coverage of the community diversity as calculated by Singleton *et al.* (2001). This suggests that the majority of the predominant populations of the turkey fecal microbial community were represented in the clones analyzed. Low G+C Gram-positive bacteria were abundant among the 16S rDNA clones (i.e., 89.3%), with sequences closely related to *Lactobacillales* (47.1%), *Bacillales* (30.9%), and *Clostridiales* (11.3%) (Fig. 1 and 2). Sequences closely related to *Lactobacillus aviarius*, a numerically dominant species in chickens (Gong *et al.*, 2007), represented 18% of the turkey fecal clone sequences. *Actinobacteria* (5.5%) and γ-Proteobacteria (4.7%) were also represented, while *Bacteroidetes* were found at considerable lower levels (i.e., 0.5%). Several 16S rRNA gene clones were closely related to potential pathogens, such as *Staphylococcus lentus* (n=2, ID 99%) and *Aerococcus viridans* (n=1, ID 99%) (Birnbaum *et al.*, 1994; Facklam *et al.*, 2003). In total, 32 genera were represented in the 16S rRNA gene clone library: *Acetanaerobacterium*, *Aerococcus*, *Ammoniphilus*, *Amphibacillus*, *Anaerofilum*, *Atopostipes*, *Bacillus*, *Bacteroides*, *Brevibacterium*, *Carnobacterium*, *Cellulomonas*, *Cerasibacillus*, *Clostridium*, *Corynebacterium*, *Escherichia*, *Enterococcus*, *Eubacterium*, *Facklamia*, *Faecalibacterium*, *Gracilibacillus*, *Halobacillus*, *Jeotgalicoccus*, *Lactobacillus*, *Listeria*, *Oceanisphaera*, *Oceanobacillus*, *Ruminococcus*, *Shigella*, *Staphylococcus*, *Subdoligranulum*, *Virgibacillus*, and *Yania*.

A total of 373 turkey fecal metagenomic fragments were analyzed (i.e., 95~678 bp in length, approximately 130 kb total). Overall, BLASTX suggested that 266 DNA sequences (71.5%) had significant similarity with sequences from NR database with E-values ≤0.001 and amino acid sequence identity ≥30%. Nearly 30% of the sequences (i.e., 106) had no significant similarity to the NR database (i.e., based on both E-values and percent of sequence identity). A small number of sequences were similar to bacteriophage proteins (mainly Gram-positive bacterial phages; 9.9%) and avian or fowl adenovirus (3.2%) (Table 1). Most identified sequences (n=216 or 58.1%) were similar to bacterial proteins from 10 bacterial orders. The majority of the bacteria-like sequences were similar to proteins from members of the following bacterial orders: *Lactobacilalles* (49.5%), *Bacillales* (10.6%), and *Clostridiales* (7.9%) (Fig. 3). Other bacterial sequences were similar to *Actinobacteria* (15.3%) and proteobacterial (10.2%) proteins.

Analysis of metagenomic sequences suggested the pres-

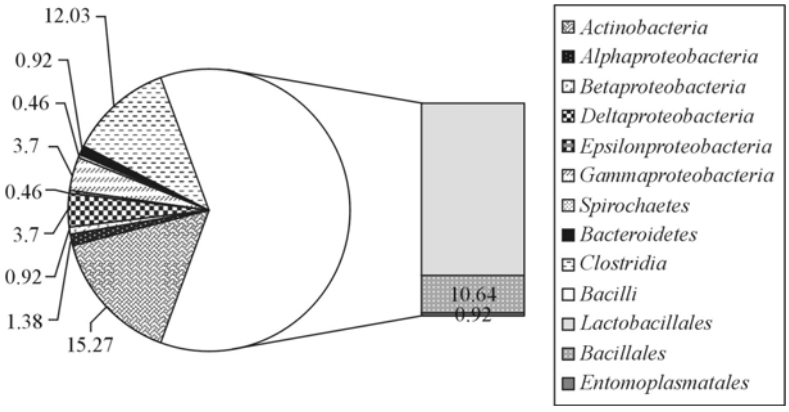

Fig. 3. Bacterial composition of turkey fecal DNA revealed by randomly amplified metagenomic sequences.

Table 2. COG functional bacterial group assignments of turkey metagenomic sequences

COG gene category	Sub COG gene category	Clone sequences (%)
Information storage and processing	[J] Translation, ribosomal structure and biogenesis	13.4
	[L] Replication, recombination and repair	10.2
	[K] Transcription	4.6
Cellular processes and signaling	[D] Cell cycle control, cell division, chromosome partitioning	0.5
	[M] Cell wall/membrane/envelope biogenesis	6.5
	[N] Cell motility	0.9
	[O] Posttranslational modification, protein turnover, chaperones	0.9
	[T] Signal transduction mechanisms	2.3
	[U] Intracellular trafficking, secretion, and vesicular transport	0.9
	[V] Defense mechanisms	0.9
Metabolism	[C] Energy production and conversion	1.9
	[E] Amino acid transport and metabolism	8.8
	[F] Nucleotide transport and metabolism	1.4
	[G] Carbohydrate transport and metabolism	9.3
	[H] Coenzyme transport and metabolism	2.3
	[I] Lipid transport and metabolism	1.9
	[P] Inorganic ion transport and metabolism	4.2
	[Q] Secondary metabolites biosynthesis, transport and catabolism	0.5
Poorly characterized	[R] General function prediction only	7.9
	[S] Function unknown	20.8

ence of bacteria involved in denitrification (*Arthrobacter* spp., *Paracoccus denitrificans*), nitrification (*Nitrosomonas europaea*, *Nitrosococcus oceani*), sulfate reduction (*Desulfitobacterium hafniense*, *Desulfuromonas acetoxidans*), and iron reduction (*Alkaliphilus metalliredigenes*). Additionally, some sequences were similar to the proteins of *Brevibacterium* (ID 32∼66%) and *Nocardioides* (ID 40∼61%) associated with the metabolism of aromatic chemicals (Iwabuchi and Harayama, 1998; Rattray and Fox, 1999), as well as *Desulfitobacterium hafniense* (ID 37%) and *Anaeromyxobacter dehalogenans* ID 41∼65%), organisms involved in dehalogenation processes (Christiansen and Ahring, 1996; Sanford *et al.*, 2002).

Function prediction for metagenomic bacteria-like sequences were categorized as information storage and processing (28.2%), cellular processes (13.3%), metabolism (30.3%), and unknown function or poorly characterized gene (28.7%) according to COG functional categories (Table 2). Of those metagenomic sequences that matched characterized functional genes, most were associated with the following subcategories: (a) translation, ribosomal structure and biogenesis (13.4%); (b) replication and recombination and repair (10.2%); (c) carbohydrate transport and metabolism (9.3%); (d) amino acid transport and metabolism (8.8%); (e) cell wall/membrane/envelope biogenesis (6.5%). Overall, the proportion of gene-function subcategories in the turkey fecal metagenomic library was similar to those of bacterial genomes in the COG database (Table 1).

Twenty eight sequences (22%) similar to metabolic functions of intestinal bacteria were identified. For example, sequences associated with *Lactobacillus*-like genes included proteins involved in the metabolism of key biomolecules like amino acids (gamma-aminobutyrate permease), nucleotides (thymidylate synthase; purine-cytosine transport protein), carbohydrates (sucrose PTS transporter, phosphoglucomutase/phosphomannomutase, glucose-6-phosphate isomerase, maltose phosphorylase, dihydroxyacetone kinase, fructose-bisphosphate aldolase), and lipids (diphosphomevalonate decarboxylase). Some sequences were similar to predicted hydrolases and lipoprotein which are involved in cellular processes and signaling, and metabolism (e.g., conjugated bile salt hydrolase beta, limonene-1,2-epoxide hydrolase catalytic domain, prolipoprotein diacylglyceryltransferase). Other sequences were similar to predicted membrane proteins, such as proteins involved in the export of O-antigen and teichoic acids, while others showed sequence similarity to transposase genes and tetracycline resistance genes.

Discussion

Limited information is currently available on the identity and microbial diversity of the turkey fecal microbial community. In our study, both 16S rRNA gene and metagenome sequence analyses showed that the adult turkey fecal bacterial community is dominated by low G+C Gram-positive bacteria. This bacterial group is also prevalent in the gastrointestinal tracts of chickens (Zhu *et al.*, 2002; Lu *et al.*, 2003). Recently, Scupham (2007a) used automated ribosomal intergenic spacer analysis and 16S rRNA gene sequencing data to study succession in the intestinal microbiota of preadolescent turkeys. *Clostridiales* were the dominant group in nine week old turkeys, while their numbers decreased considerably by week 11. During this time the percentage of *Bacteroides*-like sequences increased, suggesting changes in community structure due to age. In another study using 3-

week-old turkeys, the active cecal microbial fraction was determined by coupling bromouridine incorporation and 16S rRNA gene-based methods (Supham, 2007b). In the previous study *Clostridales* were found to be most metabolically active bacteria, while γ-Proteobacteria and *Bacteroides* were also found to be active, yet in much lower numbers. Analysis of sequencing data from our study also indicated that *Bacteroides*-like populations are consistently found in low proportions in turkey fecal samples. Similar findings have also been documented for chicken (Lu *et al.*, 2007), geese (Lu and Santo Domingo, unpublished data), and gull (Lu *et al.*, in press). The results from these studies suggest that there are similarities between the turkey fecal microbial community structure and other avian fecal microbial communities.

In addition to sequences presumed to be associated with intestinal bacteria, some turkey fecal metagenomic sequences obtained might not be of strict fecal origin. For example, sequences associated with *Actinobacteria*, including *Corynebacterium* spp., *Arthrobacter* spp., and γ-Proteobacteria, including *Acinetobacter*, are often associated with soil and plant material and not with gut systems. Sequences similar to these groups have also been obtained in relative high proportion in chicken feces (8.0% of *Actinobacteria* and 9.5% of γ-Proteobacteria; Lu *et al.*, 2007) and Canadian geese feces (2.9% of *Actinobacteria* and 54.8% of γ-Proteobacteria, unpublished data). Previous studies showed that *Corynebacterium*, *Arthrobacter*, and *Acinetobacter* are nutritionally versatile bacteria (Buchan *et al.*, 2001; Barbe *et al.*, 2004) and therefore it is not surprising to find them in the gut system. However, the ubiquity of these populations in other environments (i.e., soil) suggests that these are transient organisms rather than critical members of the turkey gut microbiota.

Metagenomic sequences closely related to genes present in pathogenic bacteria and perhaps relevant to human health were obtained in this study. For example, three sequences were 100% identical to the *Corynebacterium jeikeium* transposase gene, reported to be involved in multiple antibiotic resistance (Rosato *et al.*, 2001). These genes are important in opportunistic infections of immunocompromised patients and in the development and spread of antibiotic resistant genes in food bacteria. In addition, two sequences nearly identical (i.e., 98%) to tetW in *Bifidobacterium* sp. were part of the clone library. Interestingly, in a metagenomic analysis of chicken fecal DNA, clones 99% identical to bifidobacterial tetW were also identified (Lu *et al.*, 2007). This is not surprising since therapeutic tetracyclines are commonly used in poultry production. However, Fairchild *et al.* (2005) found that various bacteria in chicken droppings expressed tetracycline resistance, although bifidobacterial tetW were not found in their study. While the pathogenic nature of these specific populations is uncertain, it should be noted that there are concerns regarding the fact that staphylococci and enterococci of animal origin may serve as a reservoir for transmitting antibiotic resistances to human commensal flora (Joseph *et al.*, 2001).

Clones with high similarity to avian adenoviruses (DNA viruses) sequences, specifically structure-predicted proteins, were identified in this study. Analysis of chicken (Lu *et al.*, 2007) and Canadian geese (unpublished data) metagenomic data also identified sequences similar to structure proteins of avian viruses. Approximately 29% of viral sequences identified in human fecal metagenomic studies were predicted to encode for structural genes (Breitbart *et al.*, 2003). Since structure-like capsid genes (Zhong *et al.*, 2002) have been successfully used in the phylogenetic analysis of aquatic viral communities, our data further suggest that viral structural genes could be used in phylogenetic analysis of environmental viruses and to identify new viruses in fecal microbial communities.

In summary, this is the first description of the turkey fecal microbial community using a combination of 16S rRNA gene and fecal metagenome sequence analysis. As relying on sequence analysis of a single gene can introduce relevant biases, random sequencing of metagonomic fragments represents an alternate approach to those interested in describing *in toto* the community structure of natural microbial communities. In this study, metagenomic sequence analysis revealed genes that might be associated with specific microbial metabolic functions in the turkey gut. While additional studies have to be conducted using turkey intestinal samples to further confirm the predominance and potential importance of these bacterial populations to the host, it should be noted that comparative metagenomics have revealed the presence of novel gene families that perform core functions in the gut microbiota (Kurokawa *et al.*, 2007). Moreover, the approach herein described unveiled the presence of genes associated with viruses that otherwise will not be detected using bacterial based genetic assays. This information is important as it has important implications to the dynamics of gut bacteria (i.e., bacteriophage genes) and both human and animal health. As additional metagenomic information is obtained from bird fecal communities, this information will be useful in comparative metagenomic studies to those interested in understanding the genetic network and ecological roles of avian gut microbial populations and for the identification of habitat and community specific genes (Tringe *et al.*, 2005).

Acknowledgements

This research was funded in part by a New Start Award from the U.S. EPA National Center for Computational Toxicology to JSD. JL was the recipient of a National Research Council fellowship. We are grateful to Don Stoeckel for providing fecal samples. Any opinions expressed in this paper are those of the author(s) and do not, necessarily, reflect the official positions and policies of the U.S. EPA. Any mention of products or trade names does not constitute recommendation for use.

References

Acinas, S.G., L.A. Marcelino, V. Klepac-Ceraj, and M.F. Polz. 2004. Diverge and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *J. Bacteriol.* 186, 2629-2635.

Althschul, S.F., F. Thomas, L. Madden, A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.

Barbe, V., D. Vallenet, N. Fonknechten, A. Kreimeyer, S. Oztas, L. Labarre, S. Cruveiller, C. Robert, S. Duprat, P. Wincker, L.N.

Ornston, J. Weissenbach, P. Marlière, G.N. Cohen, and C. Médigue. 2004. Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. *Nucleic Acids Res.* 32, 5766-5779.

Birnbaum, D., L. Herwaldt, D.E. Low, M. Noble, M. Pfaller, R. Sherertz, and A. Chow. 1994. Efficacy of microbial identification system for epidemiologic typing of coagulase negative staphylococci. *J. Clin. Microbiol.* 32, 2113-2119.

Breitbart, M., P. Salamon, B. Andresen, J.M. Mahaffy, A.M. Segall, D. Mead, F. Azam, and F. Rohwer. 2003. Genomic analysis of uncultured marine viral communities. *J. Bacteriol.* 185, 6220-6223.

Buchan, A., E.L. Neidle, and M.A. Moran. 2001. Diversity of the ring-cleaving dioxygenase gene *pcaH* in a salt marsh bacterial community. *Appl. Environ. Microbiol.* 67, 5801-5809.

Case, R.J., Y. Boucher, I. Dahllöf, C. Holmström, W.F. Doolittle, and S. Kjelleberg. 2007. Use of 16S rRNA and *rpoB* genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol.* 73, 278-288.

Christiansen, N. and B.K. Ahring. 1996. *Desulfitobacterium hafniense* sp. nov., an anaerobic, reductively dechlorinating bacterium. *Int. J. Syst. Bacteriol.* 46, 442-448.

Cole, J.R., B. Chai, R.J. Farris, Q. Wang, A.S. Kulam-Syed-Mohideen, D.M. McGarrell, A.M. Bandela, E. Cardenas, G.M. Garrity, and J.M. Tiedje. 2007. The ribosomal database project (RDP-II): introducing *myRDP* space and quality controlled public data. *Nucleic Acids Res.* 35, D169-172.

Collins, M.D., R.A. Hutson, E. Falsen, E. Inganas, and M. Bisgaard. 2002. *Streptococcus gallinaceus* sp. nov., from chickens. *Int. J. Syst. Evol. Microbiol.* 52, 1161-1164.

Cotta, M.A., T.R. Whitehead, and R.L. Zeltwanger. 2003. Isolation, characterization and comparison of bacteria from swine feces and manure storage pits. *Environ. Microbiol.* 5, 737-745.

Crosby, L.D. and C.S. Criddle. 2003. Understanding bias in microbial community analysis techniques due to *rrn* operon copy number heterogeneity. *Biotechniques* 34, 790-794.

Daly, K., C.S. Stewart, H.J. Flint, and S.P. Shirazi-Beechey. 2001. Bacterial diversity within the equine large intestine as revealed by molecular analysis of cloned 16S rRNA genes. *FEMS Microbiol. Ecol.* 38, 141-151.

Eckburg, P.B., E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gill, K.E. Nelson, and D.A. Relman. 2005. Diversity of the human intestinal microbial flora. *Science* 308, 1635-1638.

Edwards, E.J., N.R. McEwan, A.J. Travis, and R.J. Wallace. 2004. 16S rDNA library-based analysis of ruminal bacterial diversity. *Antonie Van Leeuwenhoek* 86, 263-281.

Facklam, R., M. Lovgren, P.L. Shewmaker, and G. Tyrrell. 2003. Phenotypic description and antimicrobial susceptibilities of *Aerococcus sanguinicola* isolates from human clinical samples. *J. Clin. Microbiol.* 41, 2587-2592.

Fairchild, A.S., J.L. Smith, U. Idris, J. Lu, S. Sanchez, L.B. Purvis, C. Hofacre, and M.D. Lee. 2005. Effects of orally administered tetracycline on the intestinal community structure of chickens and on *tet* determinant carriage by commensal bacteria and *Campylobacter jejuni*. *Appl. Environ. Microbiol.* 71, 5865-5872.

Gill, S.R., M. Pop, R.T. Deboy, P.B. Eckburg, P.J. Turnbaugh, B.S. Samuel, J.I. Gordon, D.A. Relman, C.M. Fraser-Liggett, and K.E. Nelson. 2006. Metagenomic analysis of the human distal gut microbiome. *Science* 312, 1355-1359.

Gong, J., W. Si, R.J. Forster, R. Huang, H. Yu, Y. Yin, C. Yang, and Y. Han. 2007. 16S rRNA gene-based analysis of mucosa-associated bacterial community and phylogeny in the chicken gastrointestinal tracts: from crops to ceca. *FEMS Microbiol. Ecol.* 59, 147-157.

Huber, T., G. Faulkner, and P. Hugenholtz. 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* 20, 2317-2319.

Iwabuchi, T. and S. Harayama. 1998. Biochemical and molecular characterization of 1-hydroxy-2-naphthoate dioxygenase from *Nocardioides* sp. KP7. *J. Biol. Chem.* 273, 8332-8336.

Janssen, P.H. 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl. Environ. Microbiol.* 72, 1719-1728.

Joseph, S.W., J.R. Hayes, L.L. English, L.E. Carr, and D.D. Wagner. 2001. Implication of multiple antimicrobial-resistant enterococci associated with the poultry environment. *Food Addit. Contam.* 18, 1118-1123.

Kurokawa, K., T. Itoh, T. Kuwahara, K. Oshima, H. Toh, A. Toyoda, H. Takami, H. Morita, V.K. Sharma, T.P. Srivastava, T.D. Taylor, H. Noguchi, H. Mori, Y. Ogura, D.S. Ehrlich, K. Itoh, T. Takagi, Y. Sakaki, T. Hayashi, and M. Hattori. 2007. Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res.* 14, 169-181.

Leser, T.D., J.Z. Amenuvor, T.K. Jensen, R.H. Lindecrona, M. Boye, and K. Møller. 2002. Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. *Appl. Environ. Microbiol.* 68, 673-690.

Liles, M.R., B.F. Manske, S.B. Bintrim, J. Handelsman, and R.M. Goodman. 2003. A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Appl. Environ. Microbiol.* 69, 2684-2691.

Lu, J., U. Idris, B.G. Harmon, C. Hofacre, J.J. Maurer, and M.D. Lee. 2003. Diversity and succession of the intestinal bacterial community of the maturing broiler chicken. *Appl. Environ. Microbiol.* 69, 6816-6824.

Lu, J., J.W. Santo Domingo, and O.C. Shanks. 2007. Identification of chicken-specific fecal microbial sequences using a metagenomic approach. *Water Res.* 41, 3561-3574.

Ludwig, W. and K.H. Schleifer. 2000. How quantitative is quantitative PCR with respect to cell counts? *Syst. Appl. Microbiol.* 23, 556-562.

Marois, C., C. Savoye, M. Kobisch, and I. Kempf. 2002. A reverse transcription-PCR assay to detect viable *Mycoplasma synoviae* in poultry environmental samples. *Vet. Microbiol.* 89, 17-28.

Pearl, F.M.G., D. Lee, J.E. Bray, I. Sillitoe, A.E. Todd, A.P. Harrison, J.M. Thornton, and C.A. Orengo. 2000. Assigning genomic sequences to CATH. *Nucleic Acids Res.* 28, 277-282.

Rattray, F.P. and P.F. Fox. 1999. Aspects of enzymology and biochemical properties of *Brevibacterium linens* relevant to cheese ripening: a review. *J. Dairy Sci.* 82, 891-909.

Rondon, M.R., P.R. August, A.D. Bettermann, S.F. Brady, T.H. Grossman, M.R. Liles, K.A. Loiacono, B.A. Lynch, I.A. MacNeil, C. Minor, C.L. Tiong, M. Gilman, M.S. Osburne, J. Clardy, J. Handelsman, and R.M. Goodman. 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* 66, 2541-2547.

Rosato, A.E., B.S. Lee, and K.A. Nash. 2001. Inducible macrolide resistance in *Corynebacterium jeikeium*. *Antimicrob. Agents Chemother.* 45, 1982-1989.

Sanford, R.A., J.R. Cole, and J.M. Tiedje. 2002. Characterization and description of *Anaeromyxobacter dehalogenans* gen. nov., sp. nov., an aryl-halorespiring facultative anaerobic *Myxobacterium*. *Appl. Environ. Microbiol.* 68, 893-900.

Schmeisser, C., C. Stöckigt, C. Raasch, J. Wingender, K.N. Timmis, D.F. Wenderoth, H.C. Flemming, H. Liesegang, R.A. Schmitz, K-E. Jaeger, and W.R. Streit. 2003. Metagenome survey of biofilms in drinking-water networks. *Appl. Environ. Microbiol.* 69, 7298-7309.

Scupham, A.J. 2007a. Succession in the intestinal microbiota of preadolescent turkeys. *FEMS Microbiol. Ecol.* 60, 136-147.

Scupham, A.J. 2007b. Examination of the microbial ecology of the avian intestine *in vivo* using bromouridine. *Environ. Microbiol.* 9, 1801-1809.

Singleton, D.R., M.A. Furlong, S.L. Rathbun, and W.B. Whitman. 2001. Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. *Appl. Environ. Microbiol.* 67, 4374-4376.

Suau, A., R. Bonnet, M. Sutren, J.J. Godon, G.R. Gibson, M.D. Collins, and J. Doré. 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl. Environ. Microbiol.* 65, 4799-4807.

Suzuki, M.T., L.T. Taylor, and E.F. DeLong. 2000. Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl. Environ. Microbiol.* 66, 4605-4614.

Tatusov, R.L., E.V. Koonin, and D.J. Lipman. 1997. A genomic perspective on protein families. *Science* 278, 631-637.

Tringe, S.G., C. Von Mering, A. Kobayashi, A.A. Salamov, K. Chen, H.W. Chang, M. Podar, J.M. Short, E.J. Mathur, J.C. Detter, P. Bork, P. Hugenholtz, and E.M. Rubin. 2005. Comparative metagenomics of microbial communities. *Science* 308, 554-557.

Zhong, Y., F. Chen, S.W. Wilhelm, L. Poorvin, and R.E. Hodson. 2002. Phylogenetic diversity of marine cyanophage isolates and natural virus communities as revealed by sequences of viral capsid assembly protein gene g20. *Appl. Environ. Microbiol.* 68, 1576-1584.

Zhu, X.Y., T. Zhong, Y. Panya, and R.D. Joerger. 2002. 16S rRNA-based analysis of microbiota from the cecum of broiler chickens. *Appl. Environ. Microbiol.* 68, 124-137.