The ISME Journal (2011) 5, 639–649

www.nature.com/ismej

ORIGINAL ARTICLE

Phylogenetic and gene-centric metagenomics of the canine intestinal microbiome reveals similarities with humans and mice

Kelly S Swanson[1,2,3], Scot E Dowd[4], Jan S Suchodolski[5], Ingmar S Middelbos[1], Brittany M Vester[1], Kathleen A Barry[1], Karen E Nelson[6], Manolito Torralba[6], Bernard Henrissat[7], Pedro M Coutinho[7], Isaac KO Cann[1,8], Bryan A White[1,2,8] and George C Fahey Jr[1,2]

[1]*Department of Animal Sciences, University of Illinois, Urbana, IL, USA;* [2]*Division of Nutritional Sciences, University of Illinois, Urbana, IL, USA;* [3]*Department of Veterinary Clinical Medicine, University of Illinois, Urbana, IL, USA;* [4]*Research and Testing Laboratory and Medical Biofilm Research Institute, Lubbock, TX, USA;* [5]*Gastrointestinal Laboratory, Texas A&M University, College Station, TX, USA;* [6]*J. Craig Venter Institute, Rockville, MD, USA;* [7]*Architecture et Fonction des Macromolecules Biologiques, Universites Aix-Marseille I and II, Marseille, France and* [8]*The Institute for Genomic Biology, University of Illinois, Urbana, IL, USA*

This study is the first to use a metagenomics approach to characterize the phylogeny and functional capacity of the canine gastrointestinal microbiome. Six healthy adult dogs were used in a crossover design and fed a low-fiber control diet (K9C) or one containing 7.5% beet pulp (K9BP). Pooled fecal DNA samples from each treatment were subjected to 454 pyrosequencing, generating 503 280 (K9C) and 505 061 (K9BP) sequences. Dominant bacterial phyla included the Bacteroidetes/Chlorobi group and Firmicutes, both of which comprised ~35% of all sequences, followed by Proteobacteria (13–15%) and Fusobacteria (7–8%). K9C had a greater percentage of Bacteroidetes, Fusobacteria and Proteobacteria, whereas K9BP had greater proportions of the Bacteroidetes/Chlorobi group and Firmicutes. Archaea were not altered by diet and represented ~1% of all sequences. All archaea were members of Crenarchaeota and Euryarchaeota, with methanogens being the most abundant and diverse. Three fungi phylotypes were present in K9C, but none in K9BP. Less than 0.4% of sequences were of viral origin, with >99% of them associated with bacteriophages. Primary functional categories were not significantly affected by diet and were associated with carbohydrates; protein metabolism; DNA metabolism; cofactors, vitamins, prosthetic groups and pigments; amino acids and derivatives; cell wall and capsule; and virulence. Hierarchical clustering of several gastrointestinal metagenomes demonstrated phylogenetic and metabolic similarity between dogs, humans and mice. More research is required to provide deeper coverage of the canine microbiome, evaluate effects of age, genetics or environment on its composition and activity, and identify its role in gastrointestinal disease.

The ISME Journal (2011) **5,** 639–649; doi:10.1038/ismej.2010.162; published online 21 October 2010
Subject Category: microbe–microbe and microbe–host interactions
Keywords: canine gut; gastrointestinal bacteria; metagenomics; pyrosequencing

Introduction

Despite being a member of the order Carnivora, the domestic dog is omnivorous in nature and consumes a considerable amount of dietary carbohydrate, including fibrous materials, commonly present in commercial pet foods. Dogs do not rely heavily on microbial fermentation as it pertains to energy requirements, but balanced and stable microbiota are critical for maintaining gastrointestinal health. Characterizing the canine microbiome is important for several reasons. First, similar gastrointestinal anatomy and physiology, dietary patterns, metabolic processes and intestinal disease etiology make the dog an effective human model for intestinal health and disease (as reviewed by Swanson and Schook, 2006). Second, most pet dogs in developed countries are now treated as family, with many not only living in the home but also eating, sleeping and playing with their owners. This close proximity has relevance in terms of zoonotic disease. Several recent case reports have demonstrated a direct link between human illness and pet dogs (Ngaage *et al.*, 1999; Sato *et al.*, 2000).

Correspondence: GC Fahey Jr., Department of Animal Sciences, Division of Nutritional Sciences, University of Illinois, Urbana. IL 61801, USA.
E-mail: gcfahey@illinois.edu
Received 9 February 2010; revised 8 July 2010; accepted 7 August 2010; published online 21 October 2010

Bacteroidetes and Firmicutes are the predominant microbial phyla in the human gut (Eckburg *et al.*, 2005; Gill *et al.*, 2006). Knowledge of the canine gut microbiome lags behind that of humans, but has improved recently with the increased speed and reduced cost of next-generation sequencing technologies (Suchodolski *et al.*, 2009). Recent sequencing data from our laboratories suggest that Firmicutes, Bacteroidetes and Fusobacteria co-dominate the colon of healthy dogs (Suchodolski *et al.*, 2008a; Middelbos *et al.*, 2010). Suchodolski *et al.* (2008a) compared small and large intestinal populations and noted that *Clostridiales* predominated in the duodenum and jejunum, whereas *Fusobacteriales* and *Bacteroidales* were the most abundant bacterial order in the ileum and colon. *Enterobacteriales* were more commonly observed in the small intestine than in the colon, and *Lactobacillales* were commonly present in all parts of the gastrointestinal tract. Intestinal disease is often associated with alterations in small intestinal microbiota, some of which have also been identified in dogs. Results from several recent studies have identified distinct gut microbial populations in dogs with inflammatory bowel disease as compared with healthy controls (Xenoulis *et al.*, 2008; Suchodolski *et al.*, 2010; Allenspach *et al.*, 2010). Duodenal samples from dogs with inflammatory bowel disease had reduced species richness, were enriched with the *Enterobacteriaceae* family and also differed in *Clostridiaceae*, *Bacteroidetes* and *Spirochaetes* populations. These recent experiments have provided a strong foundation on which to build, although further experimentation with greater coverage is sorely needed.

Culture-independent, 16S rRNA gene-based techniques have greatly expanded our knowledge of bacterial phylogeny, but do not provide information pertaining to function. A metagenomics approach is advantageous because it provides a view of community structure (species richness and distribution), including fungi, archaeal and viral genomes, as well as functional (metabolic) potential (Hugenholtz and Tyson, 2008). This strategy will enhance our understanding of host–microbe relationships, with application to host metabolism and disease. Recent metagenome projects have revealed the functional capacity of the gastrointestinal organisms in numerous species, including humans (Kurokawa *et al.*, 2007; Turnbaugh *et al.*, 2009), rodents (Turnbaugh *et al.*, 2006, 2008), cattle (Brulc *et al.*, 2009) and poultry (Qu *et al.*, 2008). To our knowledge, however, the canine gastrointestinal metagenome has not been characterized and was the primary objective of this experiment.

Materials and methods

Animals and diets

Six healthy adult female hound-cross dogs (*Canis lupus familiaris*; Marshall Bioresources, North Rose, NY, USA) were used. All dogs were 1.7 years old (three pairs of littermates born within 5 days of each other) and had a mean body weight of 20.3 kg (individual body weight = 17.9, 18.3, 18.7, 20.0, 21.6 and 25.3 kg). The dogs were housed individually under environmentally controlled conditions (22 °C, 12-h light:12-h dark cycle) at the Small Animal Clinic of the University of Illinois, College of Veterinary Medicine. All animal care procedures have been described by Middelbos *et al.* (2010) and were approved by the University of Illinois Institutional Animal Care and Use Committee before conducting the experiment. Experimental diets were formulated to meet all nutritional recommendations for adult dogs provided by the Association of American Feed Control Officials, (2009). Primary ingredients of both diets included brewer's rice, poultry by-product meal, poultry fat, dried egg and vitamin and mineral premixes. The control diet (C) contained no supplemental dietary fiber, whereas the fiber-supplemented diet (BP) included 7.5% beet pulp in place of brewer's rice. Control and BP diets were similar in protein (29.7 vs 28.0%), fat (19.4 vs 21.0%) and ash (6.8 vs 7.1%) composition, but contained different fiber concentrations (1.4 vs 4.5% total dietary fiber). The complete list of dietary ingredients and chemical composition is presented in Middelbos *et al.* (2010).

Experimental procedures

A crossover design with two 14-day periods was used. Dogs were randomly assigned to one of two diets in the first period and received the other diet in the second period. Dogs were fed 300 g of diet once daily, which was determined to meet the metabolizable energy needs of the heaviest dog based on National Research Council, (2006) recommendations. At each feeding, uneaten food from the previous feeding was collected and weighed. A 4-day collection phase followed a 10-day diet adaptation phase, during which fresh (within 15 min of defecation) fecal samples were collected from each dog. Fresh feces were immediately flash-frozen in liquid nitrogen and stored at −80 °C until DNA extraction.

DNA extraction

Genomic DNA was extracted and isolated from fecal samples using a modification of the method of Yu and Morrison, (2004) and described by Middelbos *et al.* (2010). After extraction, DNA was quantified using a ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA). All six samples from each treatment were pooled on an equimolar basis.

Pyrosequencing and bioinformatics

Samples were subjected to pyrosequencing using a 454 Genome Sequencer using FLX titanium reagents

(Roche Applied Science, Indianapolis, IN, USA). Both unassembled reads and assembled contigs were analyzed separately. Sequences derived from pyrosequencing were quality trimmed on the basis of N50 values. For unassembled analyses, the data set was depleted of sequences <200bp. For assembled analyses, a *de novo* assembly was performed using NGen (DNAstar, Madison, WI, USA) with a minimum match percentage of 90%, a mismatch penalty of 25, a match size of 19, gap penalty of 25 and no repeat handling. Contigs from each of the data sets were uploaded to MG-RAST (Meyer *et al.*, 2008) and WebCARMA (Gerlach *et al.*, 2009) and annotated. Unassembled reads were also loaded into MG-RAST and into IMG/M ER (Markowitz *et al.*, 2008). Comparisons were made between the canine beet pulp (K9BP) and canine control (K9C) diet-assembled canine metagenome data sets (MG-RAST accession numbers 4444165 and 4444164, respectively), and contrasts were also evaluated against 5 MG-RAST public-assembled data sets that included chicken cecum A contigs (CCA 4440285), lean mouse cecum (LMC 4440463.3), obese mouse cecum (OMC 4440464), human stool metagenome (HSM 4444130) and human F1-S feces metagenome (F1S 4440939). Parameters were limited by a maximum *e*-value of 0.01, minimum percent identity of 50, minimum alignment length of 50 and raw score maximum of 0.3. Hierarchal clustering was performed using NCSS 2007 (Kaysville, Utah). Wards minimum variance clustering was performed using unscaled Manhattan distances according to the procedure described in the study by Kaufman and Rousseeuw, (1990).

Results and discussion

Although dogs normally consume low to moderate levels of dietary fiber and do not rely heavily on fermentation for meeting their energy needs, the presence of a stable gut microbial community is crucial for intestinal health. Next generation sequencing technologies have recently been used to characterize the identity and functional capacity of a variety of microbial communities, including the gastrointestinal tracts of mammalian species. To our knowledge, however, this is the first metagenomic data set generated from the canine gastrointestinal tract. Pyrosequencing generated a total of 1008341 sequences. Assembly of the 505 061 K9BP fecal metagenome sequences resulted in 422 653 sequences being assembled into 67 761 contigs with 1937 contigs >2 kb. The average contig length was 648 bp, with an average of six sequences per contig. Of the 503 280 K9C sequences, 424 522 sequences were assembled into 66 969 contigs with 2871 contigs >2 kb. The average contig length was 799 bp, with an average of six sequences per contig. The K9 metagenome projects were submitted to NCBI with accession numbers 38 653 and 38 651. The sequences were also submitted to NCBI short read archive under accession number SRA008853.1.

Phylogenetic analysis of bacteria, archaea, fungi and viruses

For K9BP, 70.34% of the sequences evaluated were matched to SEED subsystems (using an *e*-value of 1e–5). There were 47,691 hits against the non-redundant protein database. For the K9C data set, 76.46% of the sequences matched SEED subsystems, with 51 205 non-redundant hits. Table 1 provides an overview of these phylogenetic computations. Both metagenomes had a similar microbial profile when viewed at major taxonomic levels. As expected, the Bacteroidetes/Chlorobi group and Firmicutes were the predominant phyla in our canine fecal samples, both of which represented ~35% of all sequences. Proteobacteria (13–15%) and Fusobacteria (7–8%) were the other predominant phyla present in our samples. WebCARMA analyses demonstrated similar trends in K9C (Supplementary Figure 1) and K9BP (Supplementary Figure 2) samples, with Bacteroidetes, Firmicutes, Proteobacteria, Fusobacteria and Actinobacteria being the predominant phyla present (in descending order).

16S rRNA gene-based (variable region three) pyrosequencing data from these same fecal samples highlighted the predominance of these similar phyla, but with a lower prevalence of Bacteroidetes (27–34%), Firmicutes (17–27%) and Proteobacteria (5–7%), and much higher prevalence of Fusobacteria (27–44%; Middelbos *et al.*, 2010). Although it is difficult to identify the source of variation between these methodologies, especially in Fusobacteria, biases involved with the generation of amplicons (for example, primer bias and so on) for the 16S rRNA gene-based method may have contributed to this discrepancy. In another recent study, in which near full-length 16S rRNA gene sequences were generated from a clone library of healthy canine colon samples, Firmicutes represented ~40% of sequences, whereas Bacteroidetes and Fusobacteria both represented 30% each (Suchodolski *et al.*, 2008a). These recent studies suggest that the predominant phyla of the canine gut microbiome are similar to that of humans (Eckburg *et al.*, 2005; Li *et al.*, 2008) and mice (Ley *et al.*, 2005), both of which are also dominated by Firmicutes and Bacteroidetes. The significance of an enriched Fusobacteria

Table 1 Protein hits for canine metagenomes in relation to phylogeny

	K9BP	*K9C*
Archaea	1.12%	1.09%
Bacteria	98.08%	98.24%
Eukaryota	0.40%	0.37%
Virus	0.38%	0.29%

population in the dog gut is unknown, however, and deserves more attention in future studies.

Assuming that metagenome sequencing represented a random sample of read counts and that each group was independent of the other, it seems that K9C dogs had a greater percentage of Bacteroidetes, Fusobacteria, Proteobacteria, whereas K9BP had greater read counts of Bacteroidetes/Chlorobi and Firmicutes (Table 2). Increased Firmicutes read counts in the current (metagenomic) data set were primarily due to decreased Clostridia in K9BP as compared with K9C, whereas changes in Proteobacteria were primarily related to Gammaproteobacteria. Despite the differences in overall read counts between the metagenomic and 16S rRNA gene-derived data sets, effects of dietary fiber on phylum prevalence were similar between them (Middelbos *et al.*, 2010). Microbial phylogeny is known to respond to dietary alterations, including the amount and type of dietary fiber or other bioactive food components, as demonstrated by recent experiments performed in mice (Turnbaugh *et al.*, 2008), rats (Abnous *et al.*, 2009) and humans (Li *et al.*, 2009b). However, characterizing dietary-induced changes in the canine microbiome using high-throughput sequencing technologies has not been performed until now, and deserves more attention in future experiments.

Results of this study were compared with data sets within MG-RAST that had similar data characteristics, including sequence number, longest sequence and average sequence length (Table 3). Paired data from studies were chosen, such as lean (LMC) and obese (OMC) mouse cecal metagenomes (Turnbaugh *et al.*, 2006) and two human fecal metagenomes (F1S; HSM). F1S was considered to be a healthy human fecal metagenome (Kurokawa *et al.*, 2007), whereas HSM was defined as human feces from a malnourished subject, as well as a chicken cecal metagenome (CCA) (Qu *et al.*, 2008). The results of such comparisons have inherent assumptions on the basis of the methods and version of the database used for generating the data and parameters used to assemble the data. Results were evaluated at the phylogenetic level (Figure 1) and at the metabolic level (Figure 2). Within the phylogenetic comparison based on a double hierarchical dendogram, the two canine samples clustered together and more closely with the healthy human and mouse metagenomes than with the obese mouse metagenome. The chicken cecum metagenome was the greatest outlier as might be expected. In all samples, the Bacteroidetes/Chlorobi group, Firmicutes and Proteobacteria were most abundant. The canine metagenome was most distinguished by its greater prevalence of Fusobacteria as compared with humans and mice, as can be seen in the heat map in Figure 1. The heat map also demonstrates that the canine metagenome contained lower Actinobacteria and greater Fibrobacteres/Acidobacteria as compared with humans and mice. Finally, Spirochaetes were identified in all of the metagenomes, and Firmicutes were notably lower in the chicken cecal metagenome that was largely predominated by Bacteroidetes/Chlorobi. Similar clustering against mouse and human data was performed in IMG/M ER (http://merced.jgi-psf.org/cgi-bin/mer/main.cgi). As Supplementary Figure 3 demonstrates, canine samples clustered together and were most

Table 2 Bacterial phylum profiles for the two canine metagenome samples

	K9BP	*K9C*
Bacteroidetes/chlorobi group	37.67% (18948)	36.75% (17188)
Firmicutes	34.72% (17467)	30.52% (14276)
Proteobacteria	13.08% (6579)	15.26% (7140)
Fusobacteria	7.13% (3585)	8.64% (4039)
Bacteroidetes	3.14% (1577)	4.47% (2092)
Actinobacteria	1.01% (510)	1.00% (468)
Synergistetes	0.73% (365)	0.76% (356)
Thermotogae	0.54% (270)	0.52% (241)
Spirochaetes	0.49% (247)	0.53% (250)
Cyanobacteria	0.47% (235)	0.52% (244)
Chloroflexi	0.31% (156)	0.29% (134)
Chlamydiae/verrucomicrobia group	0.26% (132)	0.30% (139)
Fibrobacteres/acidobacteria group	0.13% (67)	0.18% (83)
Planctomycetes	0.10% (48)	0.08% (36)
Deinococcus-Thermus	0.08% (41)	0.05% (25)
Aquificae	0.05% (23)	0.03% (16)
Chlorobi	0.03% (13)	0.03% (13)
Unclassified	0.07% (36)	0.07% (36)
Environmental samples	0.01% (4)	0.01% (4)

Table 3 Overview of the MG-RAST metagenomes chosen for comparison

Metagenome (MG-RAST accession)	*Matches (no.)*	*No. of SEED categories*	*Sequence number*	*Total size MB*	*Shortest*	*Longest*	*Average*
LMC (4440463.3)	4007	255	10845	8.4	77	1307	781.8
HSM (4444130.3)	52055	716	108486	74.2	93	160132	684
CCA (4440285.3)	1547	76	27476	3.3	77	2739	123.12
K9C (4444164.3)	31823	671	66969	53.2	44	36188	794
F1S (4440939.3)	13123	367	28900	38	92	16490	1315
K9BP (4444165.3)	29093	693	67761	43.6	41	14401	642
OMC (4440464.3)	3460	266	11857	9.1	112	1187	764.7

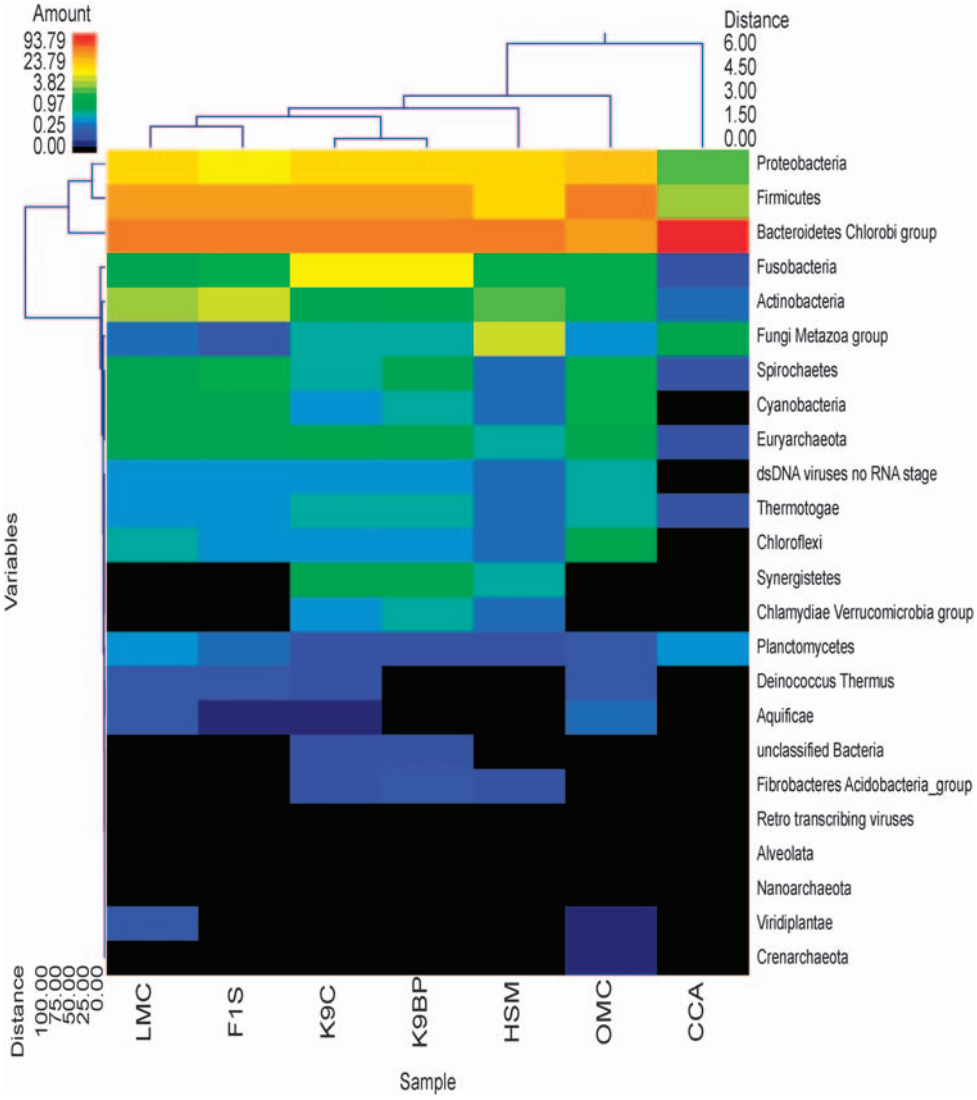

Figure 1 Phylogenetic clustering of canine, human, mouse and chicken gastrointestinal metagenomes. A double hierarchical dendogram, using the weighted-pair group clustering method and the Manhattan distance method with no scaling, shows phylogenetic distribution of microorganisms among canine (K9C; K9BP), human (F1S; HSM), murine (LMC; OMC) and chicken (CCA) metagenomes. Dendogram linkages of the bacterial classes are not phylogenetic, but based on relative abundance of the taxonomic designations within samples. The heat map depicts the relative percentage of each class of microorganism (variables clustering on the *y* axis) within each sample (*x* axis clustering). The heat map colors represent the relative percentage of the microbial designations within each sample, with the legend indicated at the upper left corner. The samples along the *x* axis with Manhattan distances are indicated by branch length and an associated scale located at the upper right corner. Clustering based on Manhattan distance of the bacterial classes along the *y* axis and their associated scale is indicated in the lower left corner.

similar to human gut samples, which also clustered together, and were followed by lean and obese mouse gut samples.

Archaea constituted a minor part of the canine metagenome, representing ~1% of all sequencing reads. No significant effect of diet on the distribution of archaea was observed. Independent of diet, two distinct archaeal phyla were identified. In the canine samples, Crenarchaeota and Euryarchaeota comprised 9 classes and 10 orders (Table 4). Methanogenic archaea were the most abundant and diverse group.

The overall abundance of archaea within the human and mice metagenome was similar to that of dogs. Cluster analysis based on the cosine similarity coefficient revealed that the sample from HSM showed the highest similarity to the canine samples (Supplemental Figure 4). Sample HSM had a high diversity of methanogens, comprising several classes (Methanobacteria, Methanococci, Methanomicrobia and Methanopyri). In contrast, human sample F1S clustered together with the mice metagenomes OMC and LMC. Although methanogens were also the most abundant group in these latter samples, they showed a lower diversity. In CCA, only one highly predominant operational taxonomic unit was evident, and this metagenome clustered separately. Compared with human, mice

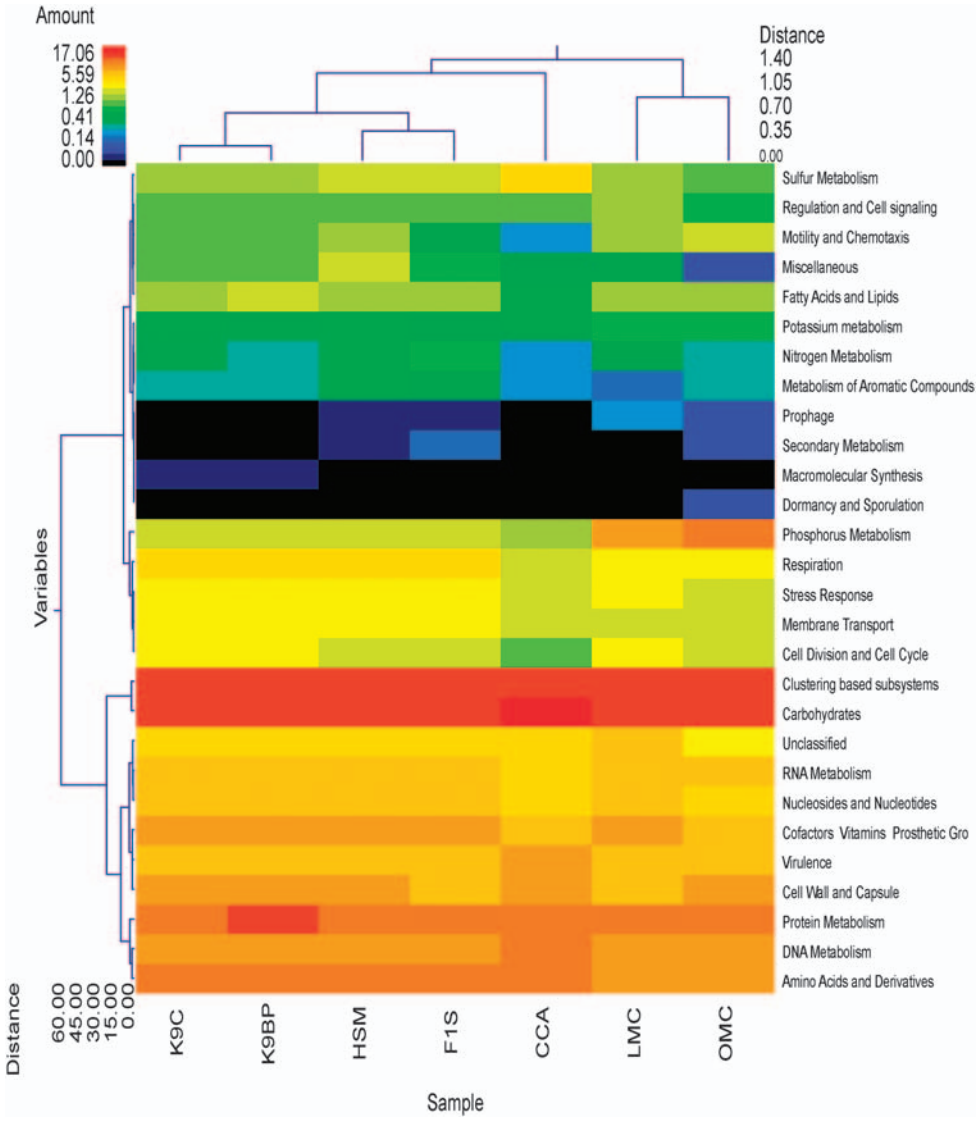

Figure 2 Metabolic clustering of canine, human, mouse and chicken gastrointestinal metagenomes. A double hierarchical dendogram, using the weighted-pair group clustering method and the Manhattan distance method with no scaling, shows bacteria distribution (classes) among canine (K9C; K9BP), human (F1S; HSM), murine (LMC; OMC) and chicken (CCA) metagenomes. Dendogram linkages are based on relative abundance of the metabolic classes (variables) within the samples. Clustering of the samples was similarly based on comparative abundance of the metabolic classes among individual samples. The heat map depicts the relative percentage of each metabolic class (variables clustering on *y* axis) within each sample (*x* axis clustering). The heat map colors represent the relative percentage of the metabolic classes within each sample, with the legend indicated at the upper left corner. The samples along the *x* axis with Manhattan distances are indicated by branch length and an associated scale located at the upper right corner. Clustering based on Manhattan distance of the metabolic classes along the *y* axis and their associated scale is indicated in the lower left corner.

and chicken metagenomes, dogs were depleted of Sulfolobales, Halobacteriales and Nanoarchaeum.

Archaea are commensal organisms in the intestine of ruminants and have also been described recently in the intestine of humans, with Methanobacteriales most commonly reported (Eckburg *et al.*, 2005; Zhang *et al.*, 2009). To our knowledge, archaea have not been described in detail in dogs. Middelbos *et al.* (2008) demonstrated *in vitro* methanogenesis in canine fecal samples, but a phylogenetic characterization of the archaeal phylotypes has not been provided. There is no clear role for archaea in the intestine. Typically, they are considered commensals, but because of mutualistic interactions with other microorganisms they may contribute to pathogenicity (Conway de Macario and Macario, 2009). Methanogens reduce hydrogen into methane, promoting an environment that enhances growth of polysaccharide fermenting bacteria, leading to a higher energy utilization of the diet. Higher numbers of methanogenic archaea have been observed in obese humans (Zhang *et al.*, 2009). Methanogens have also been associated with periodontal disease in humans (Li *et al.*, 2009a). The true prevalence and medical importance of archaea will need to be determined in dogs.

Table 4 Phylogenetic classification of Archaea for the two canine metagenome samples

Phylum	*Class*	*Order*	*Organism*	*K9BP (%)*	*K9C (%)*
Crenarchaeota	Thermoprotei	Desulfurococcales	*Aeropyrum pernix* K1	0.01	0
Euryarchaeota	Archaeoglobi	Archaeoglobales	*Archaeoglobus fulgidus* DSM 4304	0.01	0
Euryarchaeota	Halobacteria	Halobacteriales	*Haloarcula marismortui* ATCC 43049	0	0.01
Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanobrevibacter.*	0.07	0.07
			Methanosphaera stadtmanae DSM 3091	0.04	0.06
			Methanothermobacter thermautotrophicus str. Delta H	0.01	0
Euryarchaeota	Methanococci	Methanococcales	*Methanocaldococcus jannaschii* DSM 2661	0.01	0.02
			Methanococcus maripaludis C5	0.03	0.03
			Methanococcus maripaludis C6	0.01	0.03
			Methanococcus maripaludis S2	0.02	0.01
Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanocorpusculum.*	0.06	0.09
			Methanoculleus marisnigri JR1	0.02	0.01
			Methanoregula boonei 6A8	0.01	0
			Methanospirillum hungatei JF-1	0.02	0.03
Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanococcoides burtonii* DSM 6242	0.03	0.03
			Methanosaeta thermophila PT	0.01	0.01
			Methanosarcina acetivorans C2A	0.06	0.09
			Methanosarcina barkeri	0.01	0.03
			Methanosarcina barkeri str. fusaro	0.02	0.02
			Methanosarcina mazei Go1	0.02	0.04
Euryarchaeota	Methanopyri	Methanopyrales	*Methanopyrus kandleri* AV19	0	0.01
Euryarchaeota	Thermococci	Thermococcales	*Pyrococcus abyssi* GE5	0	0.01
			Pyrococcus furiosus DSM 3638	0	0.02
			Pyrococcus horikoshii OT3	0.01	0.01
			Thermococcus kodakarensis KOD1	0.01	0.01
Euryarchaeota	Thermoplasmata	Thermoplasmatales	*Ferroplasma acidarmanus*	0.01	0
			Thermoplasma acidophilum DSM 1728	0.01	0.01

A low abundance of fungi sequences was identified in the K9C and no sequences were observed in the K9BP data set. All fungi sequences were classified as Dikarya. Only three distinct phylotypes were identified, and all were at low (0.01%) abundance of the canine metagenome: *Gibberella zeae* PH-1, *Neurospora crassa* and *Saccharomyces cerevisiae*. Interestingly, the latter two phylotypes were the only ones identified in the OMC and CCA metagenomes, respectively. Fungi in the intestinal ecosystem have not yet been studied extensively in dogs. Using culture-dependent methods, only a few studies mention the presence of fungi in the canine gastrointestinal tract, including the stomach, ileum, colon and rectum, in ~25% of dogs (Davis *et al.*, 1977; Benno *et al.*, 1992; Mentula *et al.*, 2005). Using a panfungal PCR assay, a higher prevalence of fungi DNA (76% of dogs) was reported in the proximal small intestine in healthy dogs and in dogs with chronic enteropathies (Suchodolski *et al.*, 2008b). In that study, a total of 51 different phylotypes were identified across 135 dogs, but the species richness within individual dogs was low, with the majority of dogs harboring only one phylotype. Fungi DNA was detected at significantly higher proportion in mucosal brush samples than in luminal content.

Using oligonucleotide fingerprinting of ribosomal RNA genes (OFRG), a high fungi diversity and four fungi phyla were observed in the intestine of mice (Scupham *et al.*, 2006). In that study, fungi were observed mostly adjacent to the colonic epithelial cells and to a lesser extent in fecal material. Scanlan and Marchesi, (2008) measured human distal gut fungi using culture-dependent and -independent methods, demonstrating the biases that occur with culture methods. Cultured fungi were predominantly of *Candida* origin, whereas those identified by molecular techniques included *Saccharomyces, Gloeotinia, Penicillium, Candida* and *Galactomyces*. Similar fungi phylotypes (*Candida*, *Cladosporium, Penicillium* and *Saccharomyces*) were identified in stool samples from patients with human inflammatory bowel disease and from healthy controls (Ott *et al.*, 2008). In that study, fungi represented a low percentage (0.3%) of the overall fecal flora. Colon biopsy samples contained a much greater diversity of fungi, with nearly five times more operational taxonomic unit reported in biopsies compared with stool samples (Ott *et al.*, 2008). We observed a low number of fungi phylotypes in dogs in this study. It is possible that fungi constitute only a minor fraction of the gut metagenome and the true diversity has been underestimated in this study. Deeper sequencing efforts will likely yield a more accurate characterization of the fungi ecosystem. It is also possible that fungi are less prevalent in luminal content and fecal material, compared with mucosal samples, as previously shown in dogs, mice and humans (Scupham *et al.*, 2006; Ott *et al.*, 2008; Suchodolski *et al.*, 2008b), and have been therefore underestimated in this study.

Less than 1% of all sequences were classified as being of viral origin (Table 1). Only the orders Caudovirales and Iridoviridae were identified.

Within the Caudovirales, three families were observed (Myoviridae, Podoviridae and Siphoviridae) and all sequences were classified as bacteriophages (Supplementary Table 1). In dogs, the knowledge about the presence of viral communities in the gastrointestinal tract is limited to a few families (including rotavirus, coronavirus and parvovirus). Recent studies in humans revealed that the viral community in the gastrointestinal tract is highly diverse, with several hundred different phylotypes (Breitbart *et al.*, 2003). It is likely that a similar diverse viral community is present in dogs and cats, but their in-depth characterization will require deeper sequence coverage. Moreover, our approach allowed only the discovery of dsDNA. Future studies will need to survey for RNA viruses to complete our understanding of the viral intestinal ecosystem.

Metagenomics-based metabolic profiles

To our knowledge, this experiment was the first to use pyrosequencing and a metagenomics approach to characterize the metabolic capacity of canine gut microbiota and test the effects of supplemental dietary fiber. Approximately half (59.07% for K9C; 48.23% for K9BP) of all sequences in our data set were classified metabolically and are summarized in Table 5 (MG-RAST) and Supplementary Figures 5 and 6 (webCARMA). Beet pulp contains a mixture of fermentable and non-fermentable fibers and is commonly used by the pet food industry. Although phylogenetic changes were noted, the inclusion of 7.5% beet pulp did not greatly alter gene sequence number of any KEGG functional categories. The most represented functional categories included carbohydrates; protein metabolism; cell wall and capsule; cofactors, vitamins, prosthetic groups and pigments; DNA metabolism; RNA metabolism; amino acids and derivatives; and virulence.

Table 5 Metabolic profiles for the K9BP and K9C samples

Functional metabolic category	*K9BP (32678 sequences)*	*K9C (39556 sequences)*
Cofactors, vitamins, prosthetic groups and pigments	6.03% (1969)	5.67% (2240)
Cell wall and capsule	7.03% (2297)	7.61% (3008)
Potassium metabolism	0.46% (149)	0.60% (237)
Photosynthesis	0.00% (1)	0.00% (1)
Miscellaneous	1.21% (397)	1.23% (488)
Membrane transport	2.52% (824)	2.31% (913)
RNA metabolism	4.17% (1362)	3.95% (1560)
Protein metabolism	9.11% (2977)	8.12% (3210)
Nucleosides and nucleotides	3.76% (1229)	3.60% (1424)
Cell division and cell cycle	2.18% (711)	2.28% (900)
Motility and chemotaxis	0.99% (323)	0.97% (385)
Regulation and cell signaling	1.14% (373)	1.25% (495)
Secondary metabolism	0.02% (6)	0.01% (5)
DNA metabolism	7.35% (2401)	7.06% (2792)
Prophage	0.02% (5)	0.05% (19)
Unclassified	3.61% (1180)	3.64% (1439)
Virulence	6.19% (2022)	7.15% (2828)
Macromolecular synthesis	0.03% (9)	0.04% (15)
Nitrogen metabolism	0.23% (74)	0.27% (107)
Clustering-based subsystems	14.96% (4890)	14.84% (5865)
Dormancy and sporulation	0.01% (2)	0.00% (1)
Respiration	3.00% (980)	2.91% (1152)
Stress response	2.32% (757)	2.27% (899)
Sulfur metabolism	1.06% (347)	1.15% (455)
Metabolism of aromatic compounds	0.27% (89)	0.33% (132)
Amino acids and derivatives	6.80% (2222)	6.86% (2711)
Fatty acids and lipids	1.16% (379)	0.96% (380)
Phosphorus metabolism	1.89% (619)	1.85% (730)
Carbohydrates	12.50% (4084)	13.00% (5140)

Microbial carbohydrate metabolism seemed to be unaffected by diet in the present study, as the relative sequence abundance of genes related to carbohydrate metabolism was not greatly changed between dogs and was comparable with those in the core metagenome reported in a previous study of monozygotic human twins (~12.75% compared with ~12.00%, respectively; Turnbaugh *et al.*, 2009). However, certain subcategories of carbohydrate metabolism were affected by the addition of beet pulp to the diet. For example, genes related to L-rhamnose usage were twice as abundant in the K9C microbiome as compared with that of K9BP. Although these genes comprise a very small percentage of the total gene number identified (0.34% of sequences for K9C; 0.17% for K9BP), it suggests that the microbiota are being exposed to different carbohydrate concentrations when fed diets of different types or amounts of fiber substrates. Other genes related to carbohydrate metabolism, such as those related to mannose metabolism (0.54% for K9C; 0.58% for K9BP) and fructooligosaccharide and raffinose metabolism (0.29% for K9C; 0.21% for K9BP), did not reflect major differences between diets and may be due to very low dietary concentrations or low amounts present in the distal colon due to fermentation in the proximal bowel.

As a percentage of sequences, protein metabolism genes were enriched in K9BP. On further analysis, it seemed that K9BP had enriched protein biosynthesis and secretion. The largest changes within protein biosynthesis included regulation of transcription, tRNA aminoacylation (2.05% for K9C; 2.24% for K9BP) and universal GTPases (0.91% for K9C; 1.0% for K9BP). Because dogs fed the beet pulp diet (K9BP) likely had more available substrate for bacterial fermentation, this increase in protein biosynthesis simply may be the result of higher metabolic activity and/or growth of microbial populations present in this group. General protein secretion pathways were also enriched in dogs fed the beet pulp diet (0.4% for K9C; 0.51% for K9BP), but genes associated with amino acids and derivatives were unaffected by diet. Genes associated with protein degradation and protein folding seemed to be enriched in K9C.

Genes involved with the biosynthesis of vitamin K and the B vitamins were predominant in the cofactors, vitamins, prosthetic groups and pigments subsystem in the present study. Although

differences due to dietary groups were not apparent, the data set was enriched with genes associated with the biosynthesis of folate (0.91% for K9C; 0.94% for K9BP), coenzyme B_{12} (0.69% for K9C; 0.79% for K9BP), biotin (0.19% for K9C and K9BP), vitamin B_6 (0.18% for K9C and K9BP), thiamin (0.45% for K9C; 0.49% for K9BP), riboflavin (0.23% for K9C and 0.21% for K9BP) and menaquinones and phylloquinones (forms of vitamin K; 0.23% for K9C; 0.31% for K9BP). Both groups were also enriched in sequences associated with the biosynthesis of nicotinamide adenine dinucleotide (NAD) and NADP (0.58% for K9C; 0.62% for K9BP), both of which function as hydride acceptors and are important in biochemical redox reactions.

Genes related to microbial DNA and RNA metabolism were not different between dietary treatments (K9C vs K9BP). Nucleotide metabolism comprised ~4% of the 'core' metagenomic genes classified by Turnbaugh *et al*. (2009). In the present study, this value is similar to that noted for RNA metabolism and DNA metabolism. Given that the values were not different between metagenomic samples, it would be reasonable to expect that the microbiota are replicating at the same rate in each microbiome. In fact, DNA replication comprises the largest percentage of genes identified for DNA metabolism (2.12% for K9C; 2.20% for K9BP). Further, the genes associated with bacterial RNA polymerase were similar between groups (0.56% for K9C; 0.62% for K9BP).

Of those pertaining to virulence, there were a wide variety of genes associated with resistance to antibiotics and toxic compounds. A high sequence number for multidrug efflux pumps was noted in both groups, with a numerically greater prevalence in K9C (1.65%) as compared with K9BP (1.51%). In Gram-negative bacteria, these pumps provide protection by actively exporting antimicrobial substances, and serve as one mechanism by which these bacteria can survive in the presence of antibiotics (Poole, 2005). Several classes of multidrug efflux pumps exist, but all have great relevance to bacterial physiology, including antibiotic resistance, and are common drug targets (Vila and Martinez, 2008). This class of genes will continue to have great application not only in canine health but also in that of humans as well, and is an area that deserves more attention in the future. Other resistance-related genes having the greatest sequence number were those associated with acriflavine (0.35% for K9C; 0.38% for K9BP), cobalt-zinc-cadmium (0.52% for K9C; 0.53% for K9BP) and fluoroquinolones (0.50% for K9C; 0.47% for K9BP). Acriflavine is a common antiseptic agent and its resistance is thought to be due to an overexpression of efflux pump genes. The broad-spectrum antibiotics fluoroquinolones, including ciprofloxacin, levofloxacin and enrofloxacin, are commonly used in human and canine medicine, respectively. Sequences pertaining to Ton and Tol transport systems were also highly prevalent in both groups (1.01% for K9C; 0.97% for K9BP). The Ton system is an energy transducer and uses the transmembrane electrochemical gradient for nutrient uptake, whereas the Tol system is energy independent and seems to preserve the cellular envelope (Muller *et al*., 1993). These transport systems are involved in the uptake of numerous nutrients (for example, vitamin B_{12}; iron), and also serve as a mechanism by which bacteriocins translocate into and kill competing bacteria (Alonso *et al*., 2000). Finally, a few sequences for genes associated with iron scavenging were present, but very few pertained to adhesion or invasion mechanisms. Genes most prevalent in the cell wall and capsule subsystem were associated with biosynthesis of peptidoglycan (1.35% for K9C; 1.13% for K9BP), KDO2-Lipid A (0.65% for K9C; 0.63% for K9BP) and LOS core oligosaccharides (0.45% for K9C and K9BP). Rhamnose containing glycans (0.34% for K9C; 0.33% for K9BP) and sialic acid metabolism (0.73% for K9C; 0.69% for K9BP) also had a high prevalence.

To get a more in-depth view of the carbohydrate-related enzymes present in our data set, we subjected our samples to the carbohydrate-active enzymes database (CAZy; http://www.cazy.org) as described by Cantarel *et al*. (2009). K9C seemed to be enriched with glycoside hydrolases, glycosyl-transferases (765 vs 560 sequences), carbohydrate-binding modules (236 vs 140 sequences), carbohydrate esterases (236 vs 140 sequences) and polysaccharide lyases (101 vs 53 sequences) (Supplementary Tables 2 and 3). This response was counterintuitive to our hypothesis that beet pulp would increase carbohydrate-related enzymes and requires greater focus in the future. Nevertheless, although total sequence number was different between samples, the percentage of each gene within its gene family was similar for each. Even though the fiber concentration of the dog diet is rather low, cellulosomes would be expected to have a predominant role in cellulosic and hemicellulosic catabolism, as noted in other species. To our knowledge, the cellulosomes present in canine gut microbiota have not been studied thus far and would be a worthy focus in future studies.

The double hierarchical dendogram used to cluster metagenomes on the basis of metabolic capacity (Figure 2) demonstrates that canine and human samples were clustered together according to the host system, with chicken cecum and mice samples being least similar to dogs. Because this experiment used a crossover design in a research setting, dogs and housing environment were constant, leaving diet as the only difference between samples. Thus, it may not be surprising that the prevalence of each functional group was very similar between the two canine samples. The heat map also demonstrates that the canine metagenome clustered most closely with the two human metagenomes, followed by the chicken cecum metagenome. The two mouse samples clustered together and

were most different than canine samples. Interestingly, mouse samples had an enrichment of genes associated with phosphorus metabolism as compared with the other metagenomes. The chicken metagenome had a greater sequence number associated with sulfur metabolism, and fewer sequences associated with fatty acids and lipids, cell division and cell cycle and motility and chemotaxis.

To conclude, we present here the first metagenomics data set, including phylogeny and functional capacity, of the canine gastrointestinal microbiome. Our data demonstrate that the dominant bacterial phyla of the gut microbiome (for example, Bacteroidetes; Firmicutes) are similar to those of humans and rodent models. Archaea, fungi and viral sequences represented a minor portion of all sequences, but were present at levels similar to that of other mammalian biomes. Primary functional categories were also similar to those of other mammalian gut microbiomes and were associated with carbohydrates; protein, DNA and RNA metabolism; vitamin and cell-wall component biosynthesis; and virulence. Hierarchical clustering of gut metagenomic data from dogs, humans and mice demonstrated high phylogenetic and metabolic similarity among species. Although some alterations due to the inclusion of dietary fiber were noted, more drastic dietary changes (for example, source, type, or amount of macronutrients, including protein and dietary fiber) are likely needed to result in large effects in the canine gut metagenome. Although this experiment has provided a brief overview of the canine gut community, future studies are required to provide a deeper coverage and greater characterization of the metagenome of dogs in healthy and diseased states, of varying ages or genetic backgrounds, and/or receiving specific dietary interventions.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank Brandi Cantarel for her assistance with interpretation of the CAZy data.

References

Abnous K, Brooks SPJ, Kwan J, Matias F, Green-Johnson J, Selinger LB *et al.* (2009). Diets enriched in oat bran or wheat bran temporally and differentially alter the composition of the fecal community of rats. *J Nutr* **139**: 2024–2031.

Allenspach K, House A, Smith K, McNeill FM, Hendricks A, Elson-Riggins J *et al.* (2010). Evaluation of mucosal bacteria and histopathology, clinical disease activity and expression of Toll-like receptors in German shepherd dogs with chronic enteropathies. *Vet Microbiol*. doi: 10.1016/j.vetmic.2010.05.025.

Alonso G, Vilchez G, Lemoine VR. (2000). How bacteria protect themselves against channel-forming colicins. *Int Microbiol* **3**: 81–88.

Association of American Feed Control Officials (2009). *Official Publication of the Association of American Feed Control Officials*. Association of American Feed Control Officials, Inc: Oxford, IN, USA.

Benno Y, Nakao H, Uchida K, Mitsuoka T. (1992). Impact of the advances in age on the gastrointestinal microflora of beagle dogs. *J Vet Med Sci* **54**: 703–706.

Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P *et al.* (2003). Metagenomic analyses of an uncultured viral community from human feces. *J Bacteriol* **185**: 6220–6223.

Brulc JM, Antonopoulos DA, Berg Miller ME, Wilson MK, Yannarell C, Dinsdale EA *et al.* (2009). Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *PNAS* **106**: 1948–1953.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. (2009). The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res* **37**: D233–D238.

Conway de Macario E, Macario AJL. (2009). Methanogenic archaeal in health and disease: a novel paradigm of microbial pathogenesis. *Int J Med Microbiol* **299**: 99–108.

Davis CP, Cleven D, Balish E, Yale CE. (1977). Bacterial association in the gastrointestinal tract of beagle dogs. *Appl Environ Microbiol* **34**: 194–206.

Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M *et al.* (2005). Diversity of the human intestinal microbial flora. *Science* **308**: 1635–1638.

Gerlach W, Junemann S, Tille F, Goesmann A, Stoye J. (2009). WebCARMA: a web application for the functional and taxonomic classification of unassembled metagenomic reads. *BMC Bioinformatics* **10**: 430.

Gill SR, Pop M, DeBoy RT, Eckburg PB, Turnbaugh PJ, Samuel BS *et al.* (2006). Metagenomic analysis of the human distal gut microbiome. *Science* **312**: 1355–1359.

Hugenholtz P, Tyson GW. (2008). Metagenomics. *Nature* **455**: 481–483.

Kaufman L, Rousseeuw PJ. (1990). *Finding Groups in Data: An introduction to cluster analysis*. John Wiley: New York, NY, USA.

Kurokawa K, Itoh T, Kuwahara T, Oshima K, Toh H, Toyoda A *et al.* (2007). Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res* **14**: 169–181.

Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. (2005). Obesity alters gut microbial ecology. *PNAS* **102**: 11070–11075.

Li CL, Liu DL, Jiang YT, Zhou YB, Zhang MZ, Jiang W *et al.* (2009a). Prevalence and molecular diversity of Archaea in subgingival pockets of periodontitis patients. *Oral Microbiol Immunol* **24**: 343–346.

Li F, Hullar MAJ, Schwarz Y, Lampe JW. (2009b). Human gut bacterial communities are altered by addition of cruciferous vegetables to a controlled fruit- and vegetable-free diet. *J Nutr* **139**: 1685–1691.

Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H *et al.* (2008). Symbiotic gut microbes modulate human metabolic phenotypes. *PNAS* **105**: 2117–2122.

Markowitz VM, Ivanova NN, Szeto E, Palaniappan K, Chu K, Dalevi D *et al.* (2008). IMG/M: a data management and analysis system for metagenomes. *Nucleic Acid Res* **36**: D534–D538.

Mentula S, Harmoinen J, Heikkila M, Westermarck E, Rautio M, Huovinen P *et al.* (2005). Comparison between cultured small-intestinal and fecal microbiotas in beagle dogs. *Appl Environ Microbiol* **71**: 4169–4175.

Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M *et al.* (2008). The Metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**: 386.

Middelbos IS, Bauer LS, Fahey GC. (2008). *In vitro* evaluation of methanogenesis in the dog. *FASEB J* **22**: 444.

Middelbos IS, Vester Boler BM, Qu A, White BA, Swanson KS, Fahey Jr GC. (2010). Phylogenetic characterization of fecal microbial communities of dogs fed diets with or without supplemental dietary fiber using 454 pyrosequencing. *PLoS ONE* **5**: e9768.

Muller MM, Vianney A, Lazzaroni J-C, Webster RE, Portalier R. (1993). Membrane topology of the Escherichia coli TolR protein required for cell envelope integrity. *J Bacteriol* **175**: 6059–6061.

National Research Council (2006). *Nutrient requirements of dogs and cats*. National Academies Press: Washington, DC, USA.

Ngaage DL, Kotidis KN, Sandoe JAT, Nair RU. (1999). Do not snog the dog: infective endocarditis due to *Capnocytophaga canimorsus*. *Eur J Cardio-thoracic Surg* **16**: 362–363.

Ott SJ, Kuhbacher T, Musfeldt M, Rosenstiel P, Hellmig S, Rehman A *et al.* (2008). Fungi and inflammatory bowel diseases: alterations of composition and diversity. *Scand J Gastroenterol* **43**: 831–841.

Poole K. (2005). Efflux-mediated antimicrobial resistance. *J Antimicrob Chemother* **56**: 20–51.

Qu A, Brulc JM, Wilson MK, Law BF, Theoret JR, Joens LA *et al.* (2008). Comparative metagenomics reveals host specific metavirulomes and horizontal gene transfer elements in the chicken cecum microbiome. *PLoS ONE* **3**: e2945.

Sato Y, Mori T, Koyama T, Nagase H. (2000). *Salmonella Virchow* infection in an infant transmitted by household dogs. *J Vet Med Sci* **62**: 767–769.

Scanlan PD, Marchesi JR. (2008). Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and –independent analysis of faeces. *ISME J* **2**: 1183–1193.

Scupham AJ, Presley LL, Wei B, Bent E, Griffith N, McPherson M *et al* (2006). Abundant and diverse fungal microbiota in the murine intestine. *Appl Environ Microbiol* **72**: 793–801.

Suchodolski JS, Camacho J, Steiner JM. (2008a). Analysis of bacterial diversity in the canine duodenum, jejunum, ileum, and colon by comparative 16S rRNA gene analysis. *FEMS Microbiol Ecol* **66**: 567–578.

Suchodolski JS, Dowd SE, Westermarck E, Steiner JM, Spillman T, Wolcott RD *et al.* (2009). The effect of macrolide antibiotic tylosin on microbial diversity in the canine small intestine as demonstrated by massive parallel 16S rDNA sequencing. *BMC Microbiol* **9**: 210.

Suchodolski JS, Morris EK, Allenspach K, Jergens AE, Harmoinen JA, Westermarck E *et al* (2008b). Prevalence and identification of fungal DNA in the small intestine of healthy dogs and dogs with chronic enteropathies. *Vet Microbiol* **132**: 379–388.

Suchodolski JS, Xenoulis PG, Paddock C, Steiner JM, Jergens AE. (2010). Molecular analysis of the bacterial microbiota in duodenal biopsies from dogs with inflammatory bowel disease. *Vet Microbiol* **142**: 394–400.

Swanson KS, Schook LB. (2006). Canine nutritional model: influence of age, diet, and genetics on health and well-being. *Current Nutr Food Sci* **2**: 115–126.

Turnbaugh PJ, Backhed F, Fulton L, Gordon JI. (2008). Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* **3**: 213–223.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE *et al.* (2009). A core gut microbiome in obese and lean twins. *Nature* **457**: 480–484.

Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**: 1027–1031.

Vila J, Martinez JL. (2008). Clinical impact of the overexpression of efflux pump in nonfermentative gram-negative bacilli, development of efflux pump inhibitors. *Curr Drug Targets* **9**: 797–807.

Xenoulis PG, Palculict B, Allenspach K, Steiner JM, Van House AM, Suchodolski JS. (2008). Molecular-phylogenetic characterization of microbial communities imbalances in the small intestine of dogs with inflammatory bowel disease. *FEMS Microbiol Ecol* **66**: 579–589.

Yu Z, Morrison M. (2004). Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* **36**: 808–812.

Zhang HS, DiBaise JK, Zuccolo A, Kudrna D, Braidotti M, Yu Y *et al.* (2009). Human gut microbiota in obesity and after gastric bypass. *PNAS* **106**: 2365–2370.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)