Dog and human inflammatory bowel disease rely on overlapping yet distinct dysbiosis networks

Yoshiki Vázquez-Baeza[1], Embriette R. Hyde[2], Jan S. Suchodolski[3] & Rob Knight[1, 2].

[1]Department of Computer Science and Engineering, University of California, San Diego.
[2]Department of Pediatrics, University of California, San Diego.
[3]Gastrointestinal Laboratory, Department of Small Animal Clinical Sciences, College of Veterinary Medicine & Biomedical Sciences, Texas A&M University.

Introductory Paragraph

Inflammatory Bowel Disease (IBD) is an autoimmune condition that is difficult to diagnose, and where animal models of this disease have questionable human relevance[1]. Here we show that the dysbiosis network underlying IBD in dogs differs from that in humans, with some bacteria such as *Fusobacterium* switching roles between the two species (as *Bacteroides fragilis* switches roles between humans and mice)[2]. For example a dysbiosis index trained on humans fails when applied to dogs, but a dog-specific dysbiosis index achieves high correlations with the overall dog microbial community diversity patterns. In addition, a random forests classifier trained on dog-specific samples achieves high discriminatory power, even when using stool samples rather than the mucosal biopsies required for high discriminatory power in humans[2]. These relationships were not detected in previously published dog IBD datasets due to their limited sample size and statistical power[3]. Taken together, these results reveal the need to train host-specific dysbiosis networks and point the way towards a generalized understanding of IBD across different mammalian models.

Introduction

Dogs are commonly used as large animal models for drug discovery and safety assessment. The usefulness dogs as a model for IBD is yet unexplored, though not unreasonable, as dogs have been useful for studying spontaneously occurring disorders similar to those affecting people [4]. For example, dogs live in a close relationship with and share an environment with their owners, and are, therefore, frequently exposed to similar environmental factors, including enteropathogens and toxins. It is well recognized that dogs and humans suffer from similar spontaneous and lifestyle associated diseases such as obesity, allergies, diabetes mellitus, and cancer, and are often treated with similar antibiotics and drugs. IBD in humans is a chronic autoimmune disease of multifactorial aetiology and has limited treatment options. Similarly, canine idiopathic IBD is a commonly observed chronic inflammatory enteropathy that occurs spontaneously due to similar multifactorial aetiology, namely due to an interplay between an aberrant host immune system, genetics, environmental factors and gut microbiota [1]. Common clinical signs are vomiting, diarrhea, and weight loss. Histological evaluation of intestinal

41   biopsies reveal diffuse or multi-focal inflammatory cell infiltration (most commonly

42   lymphoplasmacytic, followed by eosinophilic and neutrophilic infiltration), with concurrent

43   changes in mucosal architecture (e.g., villus atrophy and fusion) [5]. Enteric protein loss may be

44   observed in severe cases. Also, in a subset of dogs, invasive and adherent *E. coli* have been

45   described, and these share common features with strains isolated from humans with Crohn's

46   disease [6]. Clinical signs may be controlled by single or combination therapy including dietary

47   modifications, antibiotics, and immunosuppressants. However, clinical relapse occurs

48   frequently, and life-long therapy may be needed. Previous small-scale studies have revealed

49   dysbiosis in the small and large intestine of dogs with IBD, with some changes in bacterial taxa

50   similar to those observed in humans with IBD [3]. For example, in both humans and dogs with

51   IBD, increases in Proteobacteria, specifically Enterobacteriaceae [7], and decreases in

52   Firmicutes, including *Faecalibacterium* and *Blautia* [8] have been reported. However, no detailed

53   studies comparing the changes in gut microbiota between humans and canines with IBD have

54   been reported to date. The aim of this study was to describe in detail the microbiome changes in

55   a large group of dogs with IBD, compare these to the microbiome of humans with IBD, assess

56   host similarities and differences, and to train a dysbiosis index composed of non-IBD and IBD

57   associated bacteria.

58

59   *Results and Discussion*

60

61   As expected based on results from human studies [2,9], but not previously clearly established in

62   canine studies [8,10], dog IBD cases and controls differ substantially in both microbial community

63   diversity and structure (Fig. 1). IBD dogs had a significantly lower alpha diversity (Mann-Whitney

64   test, p=0.003) compared to non-IBD affected dogs (Fig. 1A), however alpha diversity in this

65   population did not correlate with age, fat intake, weight, or protein intake (Supplementary Figure

66   1), nor with body condition scores (p > 0.05, see supplemental methods). Clear separation

67   (using PERMANOVA grouping samples by disease status, p=0.001 see supplemental methods)

68   between IBD dogs and controls was observed using unweighted UniFrac (Fig. 1B), and the

69   biplot shows the most abundant taxa driving these overall patterns. Consistent with [10]; antibiotic

70   history did not reveal differences within IBD affected dogs (PERMANOVA p=0.501, abx=35, no

71   abx=12), on the contrary a significant effect was observed in non-IBD dogs (PERMANOVA

72   p=0.01; abx=8, no abx=77). Finally when combined, the disease effect was significantly

73   stronger (PERMANOVA p=0.001) than the history of antibiotic usage (pseudo-F on IBD groups

74   1.99 and pseudo-F based on antibiotics groups 9.46; 1.99 < 9.46).

75

76   A random forests classifier [11] trained on the dog data achieved an AUC of 0.93 (Fig. 1C, see

77   discriminant OTUs in Supplementary Table 1), demonstrating excellent classification accuracy

78   compared to human stool samples where the AUC was only 0.63 using a much larger training

79   set [2], and achieved only AUC = 0.86 even using mucosal biopsies. Consequently, in dogs, but

80   not in humans, high classifier accuracy is achievable for IBD using only stool samples.

81

82   Encouraged by these results, we tested whether a dysbiosis (see supplemental methods)
83   network trained on stool samples alone could be used to interpret the pattern of IBD in dogs,
84   and whether this pattern overlapped the human network. We previously found that in humans,
85   substantially better correlation networks could be achieved from the mucosal biopsies, because
86   many taxa contributing to these networks were not seen in stool. Accordingly, we used the
87   techniques described [2] to generate correlation networks and a dysbiosis index for dog samples
88   (see supplemental methods for more information). Using the network together with the biplots
89   (Fig. 1B) we see that Gammaproteobacteria (specifically *Enterobacteriaceae*) were significantly
90   associated with IBD, whereas various Firmicutes such as *Clostridium* and *Ruminococcus* were
91   associated with non-IBD samples. When these features are compared with the discriminant
92   OTUs, obtained from the random forests classifier, we observe a general overlap of the lineages
93   (see Supplementary Table 1 and Supplementary Table 2). However we also see OTUs that are
94   not highlighted by the correlation network. Specifically Erysipelotrichaceae *Allobaculum* and
95   Lachnospiraceae *Blautia producta*, this is likely because these do not consistently co-occur with
96   or co-exclude other taxa.
97
98   The human dysbiosis index failed to negatively correlate with alpha diversity in dog samples
99   (Fig. 2 A.1 and A.2 correlation, 2 A.3 PCoA plot), while a dog-specific dysbiosis index showed a
100   statistically significant negative correlation in the same samples (Fig. 2 B.1 and 2 B.2
101   correlation, 2 C.3 PCoA plot, see for a reference the two groups in Fig. 2 D). Additionally we
102   tested the index with previous data [10], and although the sample size is limited, we observed
103   similar patterns (see Supplementary Figure 2). Similarly to humans, the dysbiosis index in dogs
104   is negatively correlated with phylogenetic diversity (r = -0.45, p < 0.001). However, the list of
105   "non-IBD" (co-occurring in non-IBD samples) and "IBD" (co-occurring in IBD samples) bacteria
106   only partially overlaps between host species (see Supplementary Table 2). Comparing the
107   correlation networks of the taxa for humans (as described in [2]) and the network generated for
108   the dog data (Fig. 2C) revealed overlapping, and discordant taxa. In particular, *Fusobacterium*
109   appears to be associated with IBD [2] and colorectal cancer [12] in humans but with non-IBD dog
110   samples. Of note, we previously observed high levels of *Fusobacterium* sp. in dogs [13] but also
111   carnivores of multiple species [14,15], and noted higher levels of *Fusobacterium* in dogs with more
112   access to the outdoors [16], which may correlate with a wide intake of other immunomodulatory
113   environmental bacteria. Given the limited adaptation (approximately 8.9 Mya when humans split
114   from Gorillas [17]) of the human lineage (historically omnivorous [18]) to a carnivorous diet, as
115   compared to the base of the Carnivora 40-45 Mya [19], it is possible that this taxon has yet to be
116   incorporated into the non-IBD portion of the network in humans. Nonetheless, it is important to
117   remember this is only a statistical association, and further research would need to be developed
118   to properly validate this. Consistent between human and canine networks and former studies
119   were findings in regards to decreased *Faecalibacterium* and increased *Escherichia coli* in IBD [8],
120   and these taxa seem to be important in this disease across animal hosts. Other taxa
121   (*Enterococcus and Allobaculum)* from the canine network that were associated with IBD or non-
122   IBD are generally consistent with previous results based on either small scale studies or
123   targeted PCR [10], but additional taxa were discovered here, such as *Butyricoccus* which was
124   associated with non-IBD dogs.
125

126 To measure the relative effect size of host-species and IBD, we combined humans and dogs
127 into a single PCoA plot, marked by clinical status (Supplementary Figure 4). We demonstrate
128 that at a microbial level the disease effect is smaller than the host effect (human vs. dog,
129 Supplementary Figure 4 A). Similarly the disease effect was weaker than the species effect
130 when analyzing PICRUSt [20] metagenome prediction data (PERMANOVA test grouping samples
131 by disease status and by host-species using a binary Jaccard matrix, p=0.001, pseudo-F by
132 disease 14.61, pseudo-F by species 52.59, Supplementary Figure 4 B), indicating that at both
133 the compositional and predicted functional level, species is a more significant influence on the
134 microbial community than disease (see supplemental methods and Supplementary Figure 3). In
135 both dogs and humans, predicted pathways were relatively similar across IBD and non-IBD
136 samples, with the most abundant pathways across both groups in both species including
137 "housekeeping" pathways such as transporters, ABC transporters, DNA repair and
138 recombination proteins, ribosome, purine metabolism, transcription factors, peptidases,
139 pyrimidine metabolism, and chromosome (Supplementary Figure 5). Although these most
140 abundant pathways were not significantly associated with health or disease, the abundance of
141 several lower abundance pathways was significantly different across IBD and non-IBD samples
142 in dogs, see Supplementary Table 3. No pathways were significantly higher in between disease
143 statuses in humans.
144
145 Taken together, these results have important implications for translational medicine and for
146 understanding IBD in dogs. Also, the major functional gene content was conserved across non-
147 IBD and IBD humans and dogs. While some significant predicted functions were identified
148 between non-IBD and IBD dogs, the human sample population did not have enough non-IBD
149 individuals analyzed for proper statistical power. This together with previous work suggests
150 similar functional changes within the microbiota of dogs and humans with IBD. Previous studies
151 have already shown that some treatment approaches to IBD that target the microbiome are
152 conserved across dogs and humans, such as antibiotics, dietary modulation, and probiotics. For
153 example, specific probiotic therapy has been shown to have similar effects on fecal and
154 mucosa-adherent microbiota and host immune response (i.e., increase of tight junction proteins,
155 increase in beneficial mucosa-adherent bacteria) in humans, dogs, and rat models of colitis [21-23].
156 Our study also revealed that the dysbiosis networks clearly differ in some key bacterial groups.
157 Better understanding of the similarities in these microbial networks and functional changes may
158 extend the abilities to test therapeutic approaches across multiple host species.
159
160 *Methods*
161
162 Naturally passed fecal samples were analyzed from 85 healthy dogs and 65 dogs with chronic
163 signs of gastrointestinal disease and confirmed inflammatory changes on histopathology. All
164 dogs participated in different clinical studies and leftover fecal samples were utilized for this
165 study. The protocol for sample collection was approved by the Clinical Research Review
166 Committee of the College of Veterinary Medicine, Texas A&M University (CRRC#09-06).
167
168 Dogs with clinical signs of chronic GI disease (i.e., vomiting, diarrhea, anorexia, weight loss,
169 etc.) were diagnosed with idiopathic IBD based on the World Small Animal Veterinary

170   Association (WSAVA) criteria: (i) chronic (i.e., > 3 weeks) GI signs; (ii) histopathologic evidence
171   of mucosal inflammation; (iii) inability to document other causes of GI inflammation; (iv)
172   inadequate response to dietary, antibiotic, and anthelmintic therapies, and (v) clinical response
173   to anti-inflammatory or immunosuppressive agents. Histological samples were obtained
174   endoscopically. The clinical status of each dog was evaluated using a published clinical canine
175   IBD activity index (CIBDAI). Within the IBD dogs, 41 dogs had histological confirmed
176   inflammation in the small intestine, 18 dogs had histological changes in both small intestine and
177   colon, and 5 dogs had only histological changes reported in the colon. Histological changes
178   were predominantly of lymphoplasmacytic infiltrates, with a subset of dogs also showing
179   eosinophilic and/or neutrophilic components. The mean (SD) clinical IBD activity index (CIBDAI)
180   for IBD dogs was 6.4 (3.1).
181
182   Dogs were excluded if they received antibiotics within the past 2 weeks of sample collection.
183   Data on antibiotic history was nevertheless collected: 34/65 dogs with IBD had no history of
184   prior antibiotics administration, while 13 dogs received antibiotics several weeks (>2) or months
185   before sample collection. The remaining 18 dogs in the IBD group had no information about
186   prior antibiotic use. In the healthy group (n=85), 76 dogs had not received any antibiotics, and 9
187   dogs had a history of antibiotic use, but not within the last 2 weeks of sample collection. No
188   technical replicates were generated in this study.
189
190   Sample and animal information (i.e., age, weight, gender, breed, duration of clinical signs,
191   histopathology, antibiotic usage) was obtained from clinical records. Also, if the owner provided
192   the information, the exact diet (trade name and manufacturer) fed at time of sample collection
193   was recorded in the clinical records, and the dietary macronutrients (protein, fat, and
194   carbohydrate content) were recorded from manufacturer's provided data on the labels.
195
196   Body weights ranged from 2.9 to 55 kg (mean 22 kg, SD: 14.9 kg), which was not significantly
197   different from (Mann Whitney test; p=0.087) the healthy dogs (range 0.9 to 50 kg; mean 20.3 kg,
198   SD: 10.7g). Mean age (SD) was 5.4 (3.07) in the IBD group, which was not significantly (Mann
199   Whitney test; p=0.311) different from healthy dogs (4.7, 3.2). There was a wide breed
200   distribution with 37 different breeds in the IBD group and 42 different breeds in the control
201   group. In the IBD group, Yorkshire terrier, German Shepherd dogs and Labrador Retrievers
202   (n=5 each) were most commonly represented.
203
204   Body condition scores (BCS) were assessed according to the WASAVA criteria. BCS is rated in
205   a 9-point scale that visually evaluates a dog's body composition. This score has been validated
206   against the standard dual-energy X-ray absorptiometry (DEXA) [24]. For this dataset, the BCS
207   was restricted to a subset of the healthy samples, therefore IBD vs. Healthy comparisons could
208   not be made in this case.

209   DNA Extraction and Sequencing

210

211   DNA isolation was performed as described by the Earth Microbiome Project Protocol (version
212   4_13) for 16S rRNA[25]. The full cohort included 192 samples, of which 15 were removed

213 because those subjects had acute hemorrhagic diarrhea, and had little clinical information
214 available. The remaining 28 samples did not recover enough sequences after quality control
215 including screening for low counts of reads per sample. All samples were sequenced using the
216 Illumina HiSeq platform (2 x 100 nucleotide sequences and an index read).

Data Processing

219 Demultiplexed and quality controlled sequences were then clustered against the Greengenes [26]
220 database (release 13_8) using the closed reference OTU picking protocol [27] as implemented in
221 QIIME [28] 1.9.0; these processing steps were performed using the default parameters. The OTU
222 table used for primary analysis was filtered to include only samples from subjects that presented
223 inflammatory bowel disease or that were healthy controls. Blank samples, and subjects with
224 diarrhea were removed. Finally the table was rarefied to normalize for sampling effort [29] at
225 15,000 sequences per sample, this cutoff was selected as having the best tradeoff between
226 sequences and samples per disease status category.

228 Samples from [10] were processed with the same pipeline as described above, except these were
229 rarefied at 4,500 sequences per sample.

231 To address the differences in sequencing length, the combination of samples with the human
232 dataset [2] was preceded by trimming the sequences to an even 100 nucleotides per sequence,
233 the rest of the pipeline was performed as described above. The same OTU picking protocol and
234 reference database were used, therefore allowing the combination of the studies using the OTU
235 tables. We only used the fecal samples from the human dataset.

Statistical analysis

238 Jupyter Notebooks [30] with the analysis of this dataset are available on this URL:
239 https://www.github.com/ElDeveloper/dogs/. Briefly, alpha and beta diversity calculations of
240 weighted and unweighted UniFrac [31] were performed using QIIME version 1.9.1. To assess
241 statistical significance in these comparisons we used PERMANOVA, grouping the samples by
242 disease status (healthy vs IBD), p=0.001 in both cases; pseudo-F statistics were 9.46 for
243 unweighted UniFrac and 39.65 for weighted UniFrac. ROC curves and feature importance were
244 calculated using Caret [32] and hack_ml (https://github.com/RNAer/hack_ml).

246 Statistical significance between diversity distributions were calculated using Mann-Whitney's
247 test, linear regressions and correlation coefficients were calculated using NumPy 1.10.4 [33] and
248 SciPy version 0.17, and visualized using Seaborn 0.7.0 [34]. Principal coordinates analysis plots
249 were created using Emperor 0.9.51-dev [35].

251 Metagenome predictions were performed using the Galaxy implementation of PiCRUST version
252 1.9.0. Significant differences in predicted KEGG Pathways between healthy and IBD groups
253 were calculated using a Kruskal-Wallis test with multiple corrections (FDR and Bonferroni
254 corrected p-values). The 16S data is comprised of 85 healthy dogs, 64 IBD-affected dogs, 29

255 healthy humans and 450 humans with IBD. On the other hand the PICRUSt predicted data is
256 comprised of 85 healthy dogs, 42 IBD-affected dogs, 23 healthy humans and 344 humans with
257 IBD.
258
259 The Jaccard distance was used to compare the PICRUSt predicted data. This metric compares
260 the samples on a presence/absence basis, and is not concerned with the similarity in the
261 abundances of the OTUs but rather with the overlap in their presence. Each value is computed
262 as one minus the ratio of shared OTUs to total OTUs present in the two samples.
263
264 To assess the quality of the predictions in the dog samples, we computed the "nearest
265 sequenced taxon index" (see Supplementary Figure S3), and verified that the samples had
266 acceptable values (below 0.15).
267
268 Dysbiosis Network
269
270 The dysbiosis network was calculated as described in [2], using CCREPE [36] and visualized using
271 Cytoscape [37]. The index is calculated by taking the log transform of the abundance of the ratio
272 of IBD-associated microbes to non-IBD associated microbes as determined by the correlation
273 network.
274
275 This network is created by first scoring the co-occurrence and co-exclusion patterns in the
276 samples. CCREPE uses a compositionally adjusted version of the checkerboard score [38]; the
277 results are filtered to remove non-statistically significant relationships, and to preserve the
278 largest connected component only. We represent the results as a graph where vertices are
279 microbes and edges are interaction types. Vertices are of one of two classes (IBD-associated
280 and Healthy-associated) as determined by the class where they were dominantly abundant, and
281 edges between vertices have a weight given by their adjusted checkerboard score (negative
282 values represent co-exclusions, and positive values represent co-occurrences). The specifics of
283 this processing are described in the supplemental Jupyter notebooks.
284
285 Accession Numbers
286
287 Raw sequences for the dog samples have been deposited to the European Nucleotide Archive
288 (ENA) at the following accession number ERP014919, equivalent processed OTU tables and
289 metadata can be accessed through Qiita (https://qiita.microbio.me) under study 833 - "Dog
290 models of inflammatory bowel disease".
291
292 The data for the human dataset[2] can be found in the ENA at the following accession numbers
293 ERP015241 and ERP015242, equivalent processed OTU tables and metadata can be accessed
294 through Qiita (https://qiita.microbio.me) under study identifiers 1939 and 1998 - "The Treatment-
295 Naive Microbiome in New-Onset Crohn`s Disease".
296
297 Data for the additional dog study[10] can be found at the Sequence Read Archive (SRA) of the
298 under accession number SRP040310.

321 *References*
322

323 **Main Text**

324 1   Jergens, A. E. & Simpson, K. W. Inflammatory bowel disease in veterinary medicine.
325   *Front Biosci (Elite Ed)* **4**, 1404-1419 (2012).
326 2   Gevers, D. *et al.* The treatment-naive microbiome in new-onset Crohn's disease. *Cell*
327   *Host Microbe* **15**, 382-392, doi:10.1016/j.chom.2014.02.005 (2014).
328 3   Honneffer, J. B., Minamoto, Y. & Suchodolski, J. S. Microbiota alterations in acute and
329   chronic gastrointestinal inflammation of cats and dogs. *World J Gastroenterol* **20**, 16489-
330   16497, doi:10.3748/wjg.v20.i44.16489 (2014).
331 4   Hayward, J. J. *et al.* Complex disease and phenotype mapping in the domestic dog. *Nat*
332   *Commun* **7**, 10460, doi:10.1038/ncomms10460 (2016).
333 5   Day, M. J. *et al.* Histopathological standards for the diagnosis of gastrointestinal
334   inflammation in endoscopic biopsy samples from the dog and cat: a report from the
335   World Small Animal Veterinary Association Gastrointestinal Standardization Group. *J*
336   *Comp Pathol* **138 Suppl 1**, S1-43, doi:10.1016/j.jcpa.2008.01.001 (2008).
337 6   Simpson, K. W. *et al.* Adherent and invasive Escherichia coli is associated with
338   granulomatous colitis in boxer dogs. *Infect. Immun.* **74**, 4778-4792,
339   doi:10.1128/IAI.00067-06 (2006).
340 7   Suchodolski, J. S., Dowd, S. E., Wilke, V., Steiner, J. M. & Jergens, A. E. 16S rRNA
341   gene pyrosequencing reveals bacterial dysbiosis in the duodenum of dogs with
342   idiopathic inflammatory bowel disease. *PloS one* **7**, e39333,
343   doi:10.1371/journal.pone.0039333 (2012).

344 8   Suchodolski, J. S. *et al.* The fecal microbiome in dogs with acute diarrhea and idiopathic
345   inflammatory bowel disease. *PloS one* **7**, e51907, doi:10.1371/journal.pone.0051907
346   (2012).
347 9   Frank, D. N. *et al.* Disease phenotype and genotype are associated with shifts in
348   intestinal-associated microbiota in inflammatory bowel diseases. *Inflamm Bowel Dis* **17**,
349   179-184, doi:10.1002/ibd.21339 (2011).
350 10   Minamoto, Y. *et al.* Alteration of the fecal microbiota and serum metabolite profiles in
351   dogs with idiopathic inflammatory bowel disease. *Gut Microbes* **6**, 33-47,
352   doi:10.1080/19490976.2014.997612 (2015).
353 11   Breiman, L. Random Forests. *Machine Learning* **45**, 5-32 (2001).
354 12   Castellarin, M. *et al.* Fusobacterium nucleatum infection is prevalent in human colorectal
355   carcinoma. *Genome Res.* **22**, 299-306, doi:10.1101/gr.126516.111 (2012).
356 13   Suchodolski, J. S., Camacho, J. & Steiner, J. M. Analysis of bacterial diversity in the
357   canine duodenum, jejunum, ileum, and colon by comparative 16S rRNA gene analysis.
358   *FEMS Microbiol. Ecol.* **66**, 567-578, doi:10.1111/j.1574-6941.2008.00521.x (2008).
359 14   Ley, R. E. *et al.* Evolution of mammals and their gut microbes. *Science* **320**, 1647-1651
360   (2008).
361 15   Muegge, B. D. *et al.* Diet drives convergence in gut microbiome functions across
362   mammalian phylogeny and within humans. *Science* **332**, 970-974,
363   doi:10.1126/science.1198719 (2011).
364 16   Song, S. J. *et al.* Cohabiting family members share microbiota with one another and with
365   their dogs. *eLife* **2**, e00458, doi:10.7554/eLife.00458 (2013).
366 17   Hedges, S. B., Marin, J., Suleski, M., Paymer, M. & Kumar, S. Tree of life reveals clock-
367   like speciation and diversification. *Mol. Biol. Evol.* **32**, 835-845,
368   doi:10.1093/molbev/msv037 (2015).
369 18   Price, S. A., Hopkins, S. S., Smith, K. K. & Roth, V. L. Tempo of trophic evolution and its
370   impact on mammalian diversification. *Proc Natl Acad Sci U S A* **109**, 7008-7012,
371   doi:10.1073/pnas.1117133109 (2012).
372 19   Wesley-Hunt, G. D. F., J. J. Phylogeny of the Carnivora: Basal relationships among the
373   carnivoramorphans, and assessment of the position of 'Miacoidea' relative to Carnivora.
374   *Journal of Systematic Palaeontology* **3**, 1-28 (2005).
375 20   Langille, M. G. *et al.* Predictive functional profiling of microbial communities using 16S
376   rRNA marker gene sequences. *Nature biotechnology* **31**, 814-821, doi:10.1038/nbt.2676
377   (2013).
378 21   Mardini, H. E. & Grigorian, A. Y. Probiotic mix VSL#3 is effective adjunctive therapy for
379   mild to moderately active ulcerative colitis: a meta-analysis. *Inflammatory bowel*
380   *diseases* **20**, 1562-1567, doi:10.1097/MIB.0000000000000084 (2014).
381 22   Rossi, G. *et al.* Comparison of Microbiological, Histological, and Immunomodulatory
382   Parameters in Response to Treatment with Either Combination Therapy with Prednisone
383   and Metronidazole or Probiotic VSL#3 Strains in Dogs with Idiopathic Inflammatory
384   Bowel Disease. *PloS one* **9**, e94699, doi:10.1371/journal.pone.0094699 (2014).
385 23   Uronis, J. M. *et al.* Gut microbial diversity is reduced by the probiotic VSL#3 and
386   correlates with decreased TNBS-induced colitis. *Inflamm Bowel Dis* **17**, 289-297,
387   doi:10.1002/ibd.21366 (2011).

388 **Methods**

389 24   Mawby, D. I. *et al.* Comparison of various methods for estimating body fat in dogs. *J Am*
390   *Anim Hosp Assoc* **40**, 109-114, doi:10.5326/0400109 (2004).
391 25   Consortium, E. M. P. *The Earth Microbiome Project Protocols*,
392   <http://www.earthmicrobiome.org/emp-standard-protocols/> (2015).

393 26 McDonald, D. *et al.* An improved Greengenes taxonomy with explicit ranks for ecological
394   and evolutionary analyses of bacteria and archaea. *ISME J* **6**, 610-618,
395   doi:10.1038/ismej.2011.139 (2012).
396 27 Edgar, R. C. Search and clustering orders of magnitude faster than BLAST.
397   *Bioinformatics* **26**, 2460-2461, doi:10.1093/bioinformatics/btq461 (2010).
398 28 Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing
399   data. *Nat. Methods* **7**, 335-336, doi:10.1038/nmeth.f.303 (2010).
400 29 Weiss, S. J. X., Z.; Amir, A.; Peddada, S.; Bittinger, K.; Gonzalez, A.; Lozupone, C.;
401   Zaneveld, J. R.; Vazquez-Baeza, Y.; Birmingham, A.; Knight, R. Effects of library size
402   variance, sparsity, and compositionality on the analysis of microbiome data. *PeerJ Pre-*
403   *print* (2015).
404 30 Perez, F. & Granger, B. E. IPython: A system for interactive scientific computing.
405   *Computing in Science & Engineering* **9**, 21-29, doi:Doi 10.1109/Mcse.2007.53 (2007).
406 31 Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial
407   communities. *Appl. Environ. Microbiol.* **71**, 8228-8235, doi:10.1128/AEM.71.12.8228-
408   8235.2005 (2005).
409 32 Kuhn, M. Building Predictive Models in R Using the caret Package. *Journal of Statistical*
410   *Software* **28**, 1-26 (2008).
411 33 van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy Array: A Structure for
412   Efficient Numerical Computation. *Computing in Science & Engineering* **13**, 22-30 (2011).
413 34 Waskom, M. Seaborn: statistical data visualization. (2012).
414 35 Vazquez-Baeza, Y., Pirrung, M., Gonzalez, A. & Knight, R. EMPeror: a tool for
415   visualizing high-throughput microbial community data. *Gigascience* **2**, 16,
416   doi:10.1186/2047-217X-2-16 (2013).
417 36 Schwager, E. e. a. Detecting statistically significant associations between sparse and
418   high dimensional compositional data. *Unpublished* (2016).
419 37 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of
420   biomolecular interaction networks. *Genome Res.* **13**, 2498-2504,
421   doi:10.1101/gr.1239303 (2003).
422 38 Stone, L. & Roberts, A. The Checkerboard Score and Species Distributions. *Oecologia*
423   **85**, 74-79, doi:Doi 10.1007/Bf00317345 (1990).

424
425 *Figures*
426
427 Figure 1. **Diversity overview and a comparison against humans**. (a) Comparison between
428 disease states and species of Faith's phylogenetic diversity (whiskers extend for 1.5 times the
429 IQR past the low and high quartiles). (b) Weighted UniFrac beta diversity biplot, for the dog
430 samples colored by disease status. (c) Performance comparison of a random forest classifier
431 trained to separate the healthy from the diseased samples in both humans and dogs.

432
433 Figure 2. **Human and dog dysbiosis index**. (a.1), (a.2) Human dysbiosis index describing the
434 dog samples grouped by disease status, (a.3) weighted UniFrac PCoA plot colored by dysbiosis
435 index. (b.1) and (b.2) Dog dysbiosis index describing the dog samples grouped by disease
436 state, (b.3) weighted UniFrac PCoA plot colored by dysbiosis index. (c) Correlation network
437 used to determine the dysbiosis index in dogs, colored by the bacteria associated with IBD and
438 non-IBD samples, co-exclusion edges are showed in dotted lines and co-occurrence edges are
439 showed in solid lines. (d) Unweighted UniFrac PCoA plot of the dog data. Figures a.1, a.2, b.1
440 and b.2 all display 95% confidence intervals.