Host-associated fungal communities are determined by host phylogeny and exhibit widespread associations with the bacterial microbiome

Xavier A. Harrison[1 #], Allan D. McDevitt[2], Jenny C. Dunn[3], Sarah Griffiths[4], Chiara Benvenuto[2], Richard Birtles[2], Jean P. Boubli[2], Kevin Bown[2], Calum Bridson[4,5], Darren Brooks[2], Samuel S. Browett[2], Ruth F. Carden[6,7], Julian Chantrey[8], Friederike Clever[4,9], Ilaria Coscia[2], Katie L. Edwards[10], Natalie Ferry[2], Ian Goodhead[2], Andrew Highlands[2], Jane Hopper[11], Joseph Jackson[2], Robert Jehle[2], Mariane da Cruz Kaizer[2], Tony King[11,12], Jessica M. D. Lea[5], Jessica L. Lenka[2], Alexandra McCubbin[13], Jack McKenzie[2], Bárbara Lins Caldas de Moraes[14], Denise B. O'Meara[15], Poppy Pescod[2], Richard F. Preziosi[4], Jennifer K. Rowntree[4], Susanne Shultz[5], Matthew J. Silk[1], Jennifer E. Stockdale[13,16], William O. C. Symondson[13], Mariana Villalba de la Pena[5], Susan L. Walker[10], Michael D. Wood[2], Rachael E. Antwis[2 #]

[#] Corresponding authors:

Dr Rachael Antwis, University of Salford, r.e.antwis@salford.ac.uk

Dr Xavier Harrison, University of Exeter, x.harrison@exeter.ac.uk

1. School of Biosciences, University of Exeter, UK. 2. School of Science, Engineering and Environment, University of Salford, UK. 3. School of Life Sciences, Joseph Banks Laboratories, University of Lincoln, UK 4. Ecology and Environment Research Centre, Department of Natural Sciences, Manchester Metropolitan University, UK. 5. School of Earth and Environmental Sciences, University of Manchester, UK. 6. School of Archaeology, University College Dublin, Ireland. 7. Wildlife Ecological & Osteological Consultancy, Wicklow, Ireland. 8. Institute of Veterinary Science, University of Liverpool, UK. 9. Smithsonian Tropical Research Institute, Ancon, Republic of Panama. 10. North of England Zoological Society, Chester Zoo, Upton-by-Chester, UK. 11. The Aspinall Foundation, Port Lympne Reserve, Hythe, Kent, UK. 12. School of Anthropology and Conservation, University of Kent. 13. School of Biosciences, University of Cardiff, UK. 14. Department of Zoology, Federal University of Pernambuco, Recife, Brazil. 15. School of Science & Computing, Waterford Institute of Technology, Ireland. 16. School of Life Sciences, University of Nottingham, UK

ABSTRACT

Interactions between hosts and their resident microbial communities are a fundamental component of fitness for both agents. Despite a recent proliferation of research on interactions between animals and their associated bacterial communities, comparative evidence from fungal communities is lacking, especially in natural populations. This disparity means knowledge of host-microbe interactions is biased towards the bacterial microbiome. Using samples from 49 species from eight metazoan classes, we demonstrate that the ecological distance between both fungal and bacterial components of the microbiome shift in tandem with host phylogenetic distance. Though so-called phylosymbiosis has been shown in bacterial communities, we extend previous knowledge by demonstrating that the magnitude of shifts in fungal and bacterial community structure across host phylogeny are correlated. These data are indicative of coordinated recruitment by hosts for specific suites of microbes, and potentially selection for bacterial-fungal interactions across a broad taxonomic range of host species. Using co-occurrence networks comprising both microbial groups, we illustrate that fungi form a critical component of microbial interaction networks, and that the strength and frequency of such interactions vary across host taxonomic groups. Collectively these data indicate fungal microbiomes may play a key role in host fitness and suggest an urgent need to study multiple agents of the animal microbiome to accurately determine the strength and ecological significance of host-microbe interactions.

INTRODUCTION

Multicellular organisms support diverse microbial communities that are critical for physiological functioning, immunity, development, evolution, behaviour, and even conservation[1–3]. Variability in host-associated microbiome composition may explain asymmetries among hosts in key traits including susceptibility to disease[4,5], fecundity[6], and resilience to environmental change[7]. Although the microbiota is a complex assemblage of bacteria, fungi, archaea, viruses and protozoa, the overwhelming majority of research published to date has focused solely on exploring interactions between the host and the bacterial component of its microbiome[8,9]. Despite fungal-bacterial interactions being relatively well documented in soils and plants[10–13], relatively few studies have examined the dynamics of non-bacterial components of the microbiome in animal hosts (but see [14–16]), especially in non-model organisms or wild systems. As such, our current understanding of the strength of host-microbe interactions is likely skewed by a bacteria-centric view of the microbiome.

The roles of animal-associated mycobiomes are currently not well understood; however, a small but growing body of research has identified the potential importance of resident fungal microbiota, termed the 'mycobiome', for host animal health. This may include diverse functions such as fat, carbon and nitrogen metabolism[17,18], degradation of cellulose and other carbohydrates[19], pathogen resistance[20], initiating immune pathways and regulating inflammatory responses[9,21], and even host dispersal[22]. Host phylogeny has repeatedly been shown to be an important predictor of bacterial microbiome structure in multiple vertebrate clades, a phenomenon known as 'phylosymbiosis'[23–27]. However, evidence of phylosymbiosis from other microbial kingdoms or domains is lacking and remains a major gap in our knowledge. Addressing this shortfall is vital as we likely underestimate the strength and importance of coevolution between animal hosts and their resident fungal communities. In addition, studying multiple microbial groups in concert will allow us to identify positive and negative covariances across microbial kingdoms in the abundance of taxonomic and/or functional groups that may reflect selection for interactions among microbes that are crucial determinants of animal health.

Here we used ITS and 16S rRNA amplicon sequencing to characterise fungal and bacterial communities of primarily gut and faecal samples from 49 host species across eight classes, including both vertebrates and invertebrates (Table 1). We predicted that both fungal and bacterial microbiomes

demonstrated strong signals of phylosymbiosis across the broad host taxonomic range we test here. Specifically, we predicted that patterns of phylosymbiosis within microbial kingdoms will also drive significant positive covariance in patterns of microbial community structure between microbial kingdoms within individual hosts, suggestive of evolutionary constraints that favour co-selection of specific bacterial and fungal communities in tandem. Additionally, we used network analysis to identify key bacteria-fungi interactions whilst quantifying variation in the frequency and strength of bacteria-fungi interaction networks across host taxonomic groups. Finally, we investigate the prediction that cross-kingdom phylosymbiosis may be partially driven by similarity in host dietary niche across the 32 bird and mammal species.

RESULTS

Alpha-diversity

Our data revealed consistent patterns in fungal and bacterial alpha-diversity across host taxonomic groups. Bacterial community alpha-diversity was generally greater than, or similar to, fungal community alpha-diversity at the host species level (Fig. 1A), although two species exhibited greater fungal diversity than bacterial (great tit, tsetse fly; Fig. 1A). Alpha-diversity measures remained relatively stable within a host species whether data were rarefied to 500, 1000, or 2500 reads (Figs. 1, S1, S2; Supplementary Material). Comparisons between microbial richness values within individuals (i.e. *relative* richness) using a binomial GLMM supported these patterns, indicating that bacterial richness was higher on average in 80% of cases [95% credible interval (CI) 0.55 - 0.95]. When conditioning on Class, samples from both Mammalia and Insecta were more likely to have higher bacterial diversity than fungal diversity (credible intervals not crossing zero on the link scale; Fig. 1B), though all classes had posterior mean values >0 (>50% on the probability scale). Indeed, Mammalia were more likely to have higher bacterial relative to fungal diversity than Aves in our study organisms (mean difference in probability 22.9% [1.6 - 45.7%). Variation among species in this model explained 19.5% [7.3 - 31.2%] of the variation in relative microbial richness. Using a bivariate model with both fungal and bacterial diversity as response variables to examine patterns of *absolute* microbial richness across host taxonomy, only Mammalia exhibited bacterial diversity that was consistently higher than fungal diversity when controlling for variation among species (mean difference in index

104 5.16; [3.33 - 6.96]; Fig. 1C). This model differs from the binomial version because it accounts for the
105 magnitude of difference in microbial richness, rather than just relative richness. There was no
106 evidence of positive covariance between fungal and bacterial richness values at the species level
107 (mean correlation 0.3, 95% credible intervals -0.55 - 0.86), suggesting that high diversity of one
108 microbial group does not necessarily reflect high diversity of the other. The bivariate model also
109 revealed that species identity explained 33.9% [22.2 – 44.2%] of variation in bacterial diversity, and
110 22.4% [9.8 – 35.5%] of variation in fungal diversity. Phylogenetic analyses supported these general
111 patterns (Fig. S2). For fungi, we detected phylogenetic signal in patterns of both Inverse Simpson
112 index (C_{mean} = 0.22, p = 0.021) and number of observed amplicon sequence variants (ASVs) (C_{mean} =
113 0.26, p = 0.016). For bacteria, phylogenetic signal was evident for number of ASVs (C_{mean} = 0.28, p =
114 0.016) but not inverse Simpson index (C_{mean} = 0.114, p = 0.100).

115

116 *Beta-diversity*

117 The five most abundant classes of fungi across all host species were Dothideomycetes,
118 Eurotiomycetes, Lecanoromycetes, Pezizomycetes and Sordariomycetes, for which Dothideomycetes
119 and Eurotiomycetes showed the most variation between host species (Fig. S3). The five most
120 abundant classes of bacteria were Actinobacteria, Alphaproteobacteria, Bacilli, Bacteroidia, and
121 Gammaproteobacteria, which all varied considerably among host species (Fig. S3).

122 Both fungal and bacterial community composition at the phylum level varied considerably in
123 concert with host class and order (Fig. 2A). However, notable trends include dominance of the fungal
124 phylum Ascomycota in most species, with the exception of the two Perissodactyla (odd-toed ungulate)
125 species, where most fungal ASVs belonged to the Neocallimastigomycota. Likewise, the bacterial
126 phylum Firmicutes dominated most mammal species, but Proteobacteria were more common on
127 average in birds and fish (Fig. 2A). Analysis of factors affecting bacterial and fungal community
128 structure using PERMANOVA on centred-log ratio (CLR) transformed ASV abundances revealed
129 significant phylogenetic effects of host class, order and species, as well as effects of sample storage
130 and library preparation protocol (Table 2; Fig. S4). For both microbial kingdoms, host species identity
131 explained more variation than host class or order, and this pattern remained when re-running the

132 models without sample preparation protocol effects, though this inflated the estimate of R^2 for all
133 taxonomic groupings (Table 2).

134

135 *Phylosymbiosis*

136 Consistent with our predictions, the similarity between the microbial communities of a given pair of
137 host species was proportional to the phylogenetic distance between them (ASV level: fungal cor. =
138 0.26; p = 0.001; bacterial cor. = 0.37; p = 0.001; Fig. 2B). Correlations for both bacterial and fungal
139 communities became stronger when aggregating microbial taxonomy to family (Fig. 2B). Critically, the
140 bacterial correlation was stronger than the fungal equivalent at both ASV (mean diff. 0.077 [0.053 -
141 0.103]) and family (mean diff. 0.10 [0.082 - 0.13]) levels (Fig. 2B), indicating stronger patterns of
142 phylosymbiosis for bacteria.

143 We detected a strong, significant correlation between fungal and bacterial community
144 structure of individual samples at the level of ASVs using Procrustes rotation (cor. = 0.29, p < 0.001;
145 Fig. 2C). Collapsing ASV taxonomy to genus, family, and order resulted in even stronger correlations
146 (cor. = 0.44, 0.48 & 0.43, respectively; all p < 0.001; Fig. 2C). These data indicate a coupling between
147 the structures of fungal and bacterial communities, whereby shifts in structure of one community
148 across the phylogeny also reflect consistent shifts in the other microbial group.

149

150 *Network analysis*

151 Analysis of correlations among fungal and bacterial abundances revealed differences in network
152 structure at both the host class (Fig. 3) and host species level (Figs. S5; S6). In particular, fungi of the
153 phylum Ascomycota appeared frequently in the putative interaction networks of birds, mammals and
154 amphibians (Fig. 3). There was also systematic variation in network structure among taxonomic
155 groups. Using the class-level network data in Fig. 3, we detected that Mammalia exhibited the fewest
156 components, fewest communities, and lowest modularity (Fig. 4A, Table S1), indicating lower overall
157 network subdivision relative to other animal classes. Models of species-level network data (Fig. S5)
158 revealed the frequency of positive co-occurrence between pairs of microbes also varied by class;
159 Mammalia exhibited the highest proportion of positive edges (Fig. 4B), being significantly greater than

160 those of birds (mean diff. 0.042 [0.017-0.067]) and amphibians (mean diff. 0.05 [0.002-0.112]).
161 Notably, invertebrates had a markedly lower proportion of positive edges compared to all other taxa
162 (Fig. 4). Class explained 93.2% [92.9-93.4%] of variation in edge sign. There was also clear variation
163 at the species level; for some host species, there were considerably more positive interactions (e.g.
164 yellowhammers, pygmy shrews, greater white-toothed shrews, wood mouse, woodpigeon, yellow-
165 necked mouse; Fig. S6). In some species, there were slightly more negative interactions than positive
166 (e.g. blackcap, goldfinch; Fig. S5; S6).

167

168 *Dietary analysis*

169 While there was a significant correlation between host-associated bacterial community composition
170 and dietary data for mammals ($r = 0.334$, $p = 0.002$), and a near-significant relationship between
171 fungal community composition and diet ($r = 0.142$, $p = 0.067$), there was no significant relationship
172 between dietary data and bacterial community composition ($r = 0.087$, $p = 0.211$) or fungal community
173 composition ($r = 0.026$, $p = 0.386$) for birds. Further, taxonomic differences in microbiome composition
174 based on differences in crude dietary patterns were not clear for either bacteria or fungi when the
175 microbiome composition was visualised at the family level (Figs. S7, S8). That said,
176 Alphaproteobacteria and Eurotiomycete fungi were notably absent from species that primarily ate
177 vegetation (i.e. grasses etc) and Neocallimastigomycete fungi were the predominant fungal class
178 associated with two out of four of these host species (Figs. S7, S8).

179

180 ## DISCUSSION

181 Our study represents the most wide-ranging evaluation of animal mycobiome composition, and its
182 covariation with the bacterial microbiome, undertaken to date. Our data provide novel evidence for
183 mycobiome phylosymbiosis in wild animals, indicative of close evolutionary coupling between hosts
184 and their resident fungal communities. Consistent with previous studies, we also find evidence of
185 phylosymbiosis in the bacterial microbiome[28], but crucially, we demonstrate strong and consistent
186 covariation between fungal and bacterial communities across host phylogeny, especially at higher
187 microbial taxonomic levels. These patterns are supported by complementary network analysis

188 illustrating frequent correlative links between fungal and bacterial taxa, whereby certain pairs of
189 microbes from different kingdoms are much more likely to co-occur in the microbiome than expected
190 by chance. Taken together, these data provide novel evidence consistent with recruitment by animal
191 hosts for specific fungal and bacterial communities, which in turn may reflect selection for interactions
192 between bacteria and fungi critical for host physiology and health.

193 We found marked variation among host species in microbial community richness and
194 composition for both bacteria and fungi. Complementary analyses using mixed models and
195 phylogenetic models both detected a signal of host phylogeny in determining fungal and bacterial
196 microbiome diversity. Though our data suggest many species support a diverse assemblage of host-
197 associated fungi, critically we show that i) bacterial diversity tends to be higher on average relative to
198 fungal diversity; and ii) there is no signal of positive covariance between fungal and bacterial richness
199 within species, suggesting more ASV-rich bacterial microbiomes are not consistently associated with
200 more ASV-rich mycobiomes. These patterns could arise because of competition for niche space
201 within the gut, where high bacterial diversity may reflect stronger competition that prevents
202 proliferation of fungal diversity. Fungi are enzymatically very active, and their crucial roles in
203 degrading organic molecules and redistributing carbon and nitrogen in soil and plant ecosystems are
204 well documented[29,30]. Understanding patterns of niche competition within and among microbial groups
205 requires that we are able to define those niches by measuring microbial gene function, and
206 quantifying degree of overlap or redundancy in functional genomic profiles across bacteria and fungi.

207 We detected strong phylosymbiosis for both fungi and bacteria across a broad host
208 phylogeny encompassing both vertebrate and invertebrate classes. This pattern was significantly
209 stronger in bacteria than for fungi. In both microbial kingdoms, the signal of phylosymbiosis
210 strengthened when aggregating microbial taxonomic assignments to family level, a phenomenon that
211 has previously been shown for bacterial communities[31]. That this pattern also occurs in fungi suggests
212 either that host recruitment is weaker at finer-scale taxonomies, or our ability to detect that signal is
213 weaker at the relatively noisy taxonomic scale of ASVs. Stronger signals of phylosymbiosis at family-
214 level taxonomies may reflect the deep evolutionary relationships between hosts and their bacterial
215 and fungal communities, as well as the propensity for microbial communities to allow closely related
216 microbes to establish whilst repelling less related organisms[32]. That is, higher-order microbial
217 taxonomy may better approximate functional guilds within the microbiome, such as the ability to

218 degrade cellulose[25,31], which are otherwise obscured by taxonomic patterns of ASVs. Resolving this
219 requires the integration of functional genomic data from the fungal and bacterial microbiota into the
220 phylogeny.

221 In addition to microbe-specific patterns of phylosymbiosis, a key novel finding of our work is
222 that fungal and bacterial community composition correlate strongly across the host phylogeny. These
223 patterns are consistent with host recruitment for particular suites of fungal and bacterial taxa, which
224 may represent bacteria-fungi metabolic interactions beneficial to the host. Bacteria-fungi interactions
225 have previously been demonstrated for a handful of host species[8,9,17,33,34], but here we show these are
226 widespread across multiple animal classes. Both bacteria and fungi have considerable enzymatic
227 properties that facilitate the liberation of nutrients for use by other microbes, thus facilitating cross-
228 kingdom colonisation[35] and promoting metabolic inter-dependencies[36–38]. These findings are
229 supported by our network analyses, which identified numerous putative interactions between bacteria
230 and fungi for many of our host species. Critically, the frequency and predicted direction of these
231 relationships varied considerably among host classes, with the mammalian network exhibiting i) a
232 lower modularity, indicating weaker clustering into fewer discrete units (both distinct components and
233 interlinked communities); and ii) a higher frequency of positive correlations between microbes
234 compared to most other classes, in particular birds and insects. Comparisons of networks are
235 challenging when they differ in size (i.e. number of nodes) and structure, and differences between
236 classes in traits like modularity will also be affected by species replication within each class. However,
237 proportional traits like interaction structure (proportion of positive interactions) are unlikely to be driven
238 solely by sample size, suggesting marked biological variation in strength of fungi-bacteria interactions
239 across the host phylogeny. These putative interaction networks provide novel candidates for further
240 investigation in controlled systems, where microbiome composition and therefore the interactions
241 among microbes can be manipulated to test the influence of such interactions on host physiology.

242 The drivers of phylosymbiosis remain unclear, even for bacterial communities; is it a
243 phylogenetic signal indicative of host-microbiome coevolution, or simply a product of “ecological
244 filtering” of the microbiome in the host organism either via extrinsic (e.g. diet, habitat) or intrinsic
245 sources (e.g. gut pH, immune system function)[26,28,39]? Our results indicate host diet may play a role in
246 determining overall fungal community composition, although the relationship is weak and only evident
247 for bacteria in mammals. These results are broadly consistent with previous work, where the influence

of diet on bacterial microbiome was most evident in mammals[25]. However, Li et al.[15] showed that the composition and diversity of both fungal and bacterial communities of faecal samples differed between phytophagous and insectivorous bats, and Heisel et al.[17] demonstrated changes in fungal community composition in mice fed a high fat diet. Our study was not designed to test for the effects of ecological variation in diet on fungal microbiome *within* a species. Nor can we discount the possibility that at finer taxonomic scales within classes, signals of the effect of *among* species variation in diet on mycobiome become stronger (e.g. [15]). Although diet is thought to be a predominant driver of bacterial microbiome composition in host organisms[25,40], there is also evidence that the signals produced from faecal and true gut samples differ; that is, data generated from faecal samples indicate diet is the predominant driver of "gut" microbiome composition, whereas data from gastrointestinal samples indicate host species is the predominant determinant[41]. Moreover, faecal samples may only represent a small proportion of the gastrointestinal microbiome[41–43]. Our data also show that sample type has a significant effect on both fungal and bacterial community composition (as well as DNA extraction method and storage method; see [44–47] for other examples of this). As such, a more thorough analysis of true gut communities is required to determine the extent to which mycobiome phylosymbiosis occurs across host taxa, and what other ecological and host-associated factors influence mycobiome composition and function.

These data provide strong evidence that both fungal symbionts and fungi-bacteria interactions are likely to be critical for host functioning and health. Within animals, the roles of host-associated fungal communities are less well understood, yet these data highlight that fungi are important components of microbiome structure that are often overlooked. Key priorities for future work are to i) understand the range of functions provided by the host mycobiome, and how these alter or complement those provided by the bacterial microbiome; and ii) characterise the functional interactions between bacteria and fungi and how they influence key host metabolic processes and life history.

METHODS

Sample collection

276 DNA was extracted from tissue or faecal samples of 49 host species using a variety of DNA extraction
277 methods (Table 1) and normalised to ~10 ng/ul. Samples were largely collated from previous studies
278 and/or those available from numerous researchers and as such, DNA extraction and storage
279 techniques were not standardised across species. We sequenced a median of 10 samples per
280 species (range of 5 to 12; Table 1).

281

282 *ITS1F-2 and 16S rRNA amplicon sequencing and pre-processing*

283 To identify fungal communities, we amplified DNA for the ITS1F-2 rRNA gene using single index
284 reverse primers and a modified protocol of Smith & Peay[48] and Nguyen et al.[49], as detailed in Griffiths
285 et al.[13] (see supplementary material for additional information). To identify bacterial communities, we
286 amplified DNA for the 16S rRNA V4 region using dual indexed forward and reverse primers according
287 to Kozich et al.[50] and Griffiths et al.[51] (see supplementary material for additional information). We
288 sequenced normalised libraries using paired-end reads (2 x 250bp) with Illumina v2 chemistry on the
289 MiSeq platform at the University of Salford. We ran the ITS rRNA library twice to increase sequencing
290 depth, and combined data within samples across these two runs in the data pre-processing stage.

291 We conducted all data processing and analysis in RStudio v1.2.1335 for R[52,53] (see
292 supplementary files for full code). We conducted amplicon sequence processing in DADA2 v1.5[54] for
293 both ITS rRNA and 16S rRNA amplicon data (see supplementary material for additional information).
294 After data processing, we obtained a median of 1425 reads per sample (range of 153 to 424,527) for
295 ITS rRNA libraries, and a median of 3273 reads per sample (range of 153 to 425,179) for 16S rRNA
296 libraries.

297

298 *Host Phylogenetic Distances*

299 As many of our host species lack genomic resources from which to construct a genome-based
300 phylogeny, we built a dated phylogeny of host species using TimeTree[55]. The phylogenetic tree
301 contained 42 species, of which 36 were directly represented in the TimeTree database. A further six
302 species had no direct match in TimeTree and so we used a congener as a substitute (*Amietia,*
303 *Glossina, Portunus, Ircinia, Amblyomma, Cinachyrella*). We calculated patristic distance among

304 species based on shared branch length in the phylogeny using the 'cophenetic' function in the *ape*
305 package[56] in R. We visualised and annotated the phylogeny using the R package *ggtree*[57]. To create
306 a phylogeny for all samples, we grafted sample-level tips onto the species phylogeny with negligible
307 branch lengths following Youngblut et al.[25].

308

309 *Fungal and bacterial community analysis*

310 We used the R package *brms*[58,59] to fit the following (generalized) linear mixed effects models
311 [(G)LMMs]. For higher order taxonomic predictors and random effects, we binned all invertebrate
312 classes into a single grouping to improve model performance, as otherwise invertebrate class and
313 species were co-linear. All vertebrate taxonomic groupings were equivalent to class (Mammalia, Aves
314 etc). To compare alpha-diversity between species and microbial kingdoms, we rarefied libraries to
315 500 reads per sample, yielding 292 samples from 46 species and 307 samples from 47 species for
316 fungal and bacterial kingdoms respectively. Alpha-diversity measures remained relatively stable within
317 a host species whether data were rarefied to 500, 1000, or 2500 reads, with similar patterns exhibited
318 between the two kingdoms (Figs. 1, S1, S2; see Supplementary Material for more details). To
319 visualise differences between microbial richness within species, we filtered the data to species with at
320 least two samples per microbial kingdom, giving a total of 41 species from six classes. For model
321 fitting, we filtered the data to only those samples with paired metrics of microbial richness for both
322 kingdoms (201 observations from 42 species). We fitted two models to these data. First, to quantify
323 relative differences in richness between bacteria and fungi within a sample, we used GLMMs in the
324 *brms* package, with i) Bernoulli errors and a logit link; ii) a binary response of '1' if bacterial richness
325 was higher than fungal richness, and '0' otherwise; and iii) 'Species' nested within 'Class' as random
326 intercepts. We did not include intermediate levels of taxonomy because replication at Order and
327 Family levels was low relative to Class. We did not use a phylogenetic mixed model as not all species
328 were represented in the TimeTree phylogeny. Second, to quantify absolute differences in microbial
329 richness, we fitted a bivariate response LMM with both fungal and bacterial richness values as a two-
330 column response with Class as a fixed effect, and Species as a random intercept. For all models, we
331 used uninformative Cauchy priors for the random effects and Gaussian priors for fixed effects
332 coefficients. We assessed model adequacy using visual inspection of chains to assess mixing and

stationarity properties, as well as posterior predictive checks using the 'pp_check' function in *brms*. We calculated r^2 of models using the 'bayes_R2' function. We assessed the importance of terms based on whether 95% credible intervals of the parameter estimates of interest crossed zero. We used *ggplot*[60], *cowplot*[61] and *tidybayes*[62] for raw data and model estimate plotting. To support these analyses, we also used the R packages *phylobase*[63] and *phylosignal*[64] to estimate the phylogenetic signal in patterns of alpha diversity for both bacteria and fungi, using both Inverse Simpson Index and number of observed ASVs as outcome variables. We calculated Abouheif's C_{mean} for each diversity-microbe combination and corrected p values for multiple testing using Benjamini-Hochberg correction.

To identify taxonomic differences in microbiome and mycobiome composition between host species, we used centred-log-ratio (CLR) transformation in the *microbiome*[65] package to normalise microbial abundance data, which obviates the need to lose data through rarefying[66]. To visualise differences in microbial community structure among samples, we i) plotted proportional abundance of microbial groups at the phylum level, aligned to the host phylogenetic tree, ii) agglomerated the data to class level and visualised the variation in CLR-transformed ratios for the five most abundant microbial classes in each kingdom for each species using jitter plots, and iii) conducted principal components analysis (PCA) using CLR-transformed abundance matrices for each kingdom. To quantify differences in beta-diversity among kingdoms and species whilst simultaneously accounting for sample storage and library preparation differences among samples, we conducted a PERMANOVA analysis on among-sample Euclidean distances of CLR-transformed abundances using the adonis function in *vegan*[67] with 999 permutations. For both kingdoms, we specified effects in the following order: sample type, tissue storage, extraction kit, Class, Order, Species. This marginalises the effects of sample metadata variables first, before partitioning the remaining variance into that accounted for by host phylogeny. The results were similar when amplicon data were converted to relative abundance or rarefied to 500 reads (data not presented).

To test the hypothesis that inter-individual differences in microbial community composition were preserved between microbial kingdoms, we performed Procrustes rotation of the two PCA ordinations for bacterial and fungal abundance matrices, respectively (n = 277 paired samples from 46 species). We also repeated this analysis with ASVs agglomerated into progressively higher taxonomic rankings from genus to order (see [31]). To provide a formal test of differences in strength of correlation at different taxonomic levels, we conducted a bootstrap resampling analysis where for

each kingdom at each iteration, we randomly sampled 90% of the data and recalculated the correlation metric. We repeated this process 999 times to build a distribution of correlation values at each taxonomic grouping.

To examine the hypothesis that inter-individual distance in microbial community composition varies in concert with inter-host phylogenetic distance, we performed a Procrustes rotation on the paired matrix of microbial distance (Euclidean distance of CLR-transformed abundances) and patristic distance from the phylogenetic tree. We also repeated the analysis but binning the microbial data to family level. As above, we conducted a bootstrap resampling procedure, selecting 90% of the data at random and recalculating the correlation, for a total of 1000 permutations. This allowed us to test for significant differences in strength of correlation *within* kingdoms across taxonomic grouping levels, and *across* kingdoms within a particular taxonomic grouping.

To determine the effect of diet on bacterial and fungal community composition, we used only samples from the bird and mammal species and agglomerated the data for each host species using the merge_samples function in *phyloseq*[68]. This gave us an representative microbiome for each host species, which we rarefied to the lowest number of reads for each combination of kingdom and host taxon (2,916 – 9,160 reads; bacterial read counts were low for lesser horseshoe bats and so this species was removed from this analysis) and extracted Euclidean distance matrices for each. We then correlated these with dietary data obtained from the EltonTraits database[69] using Mantel tests with Kendall rank correlations in the *vegan* package[67]. We agglomerated the microbial data to class level and visualised the bacterial and fungal community compositions for mammals alongside pie charts displaying EltonTrait dietary data for each species.

To identify potential relationships between fungal and bacterial communities, we conducted two analyses; 1) We used the R package *SpiecEasi*[70] to identify correlations between unrarefied, CLR-transformed ASVs abundances at the host class level (with Invertebrates grouped), and 2) we used co-occurrence analysis at the species level, by rarefying the bacterial and fungal data sets to 500 reads each, and agglomerated these to family level, resulting in 117 bacterial groups and 110 fungal groups. We then merged the phyloseq objects for bacterial and fungal communities for each sample, with sufficient data retained to conduct the co-occurrence analysis for 40 host species. Using these cross-kingdom data, we calculated the co-occurrence between each pair of microbial genera by

392 constructing a Spearman's correlation coefficient matrix in the *bioDist* package[71,72]. We visualised
393 those with rho > 0.50 (strong positive interactions) and rho < -0.50 (strong negative interactions) for
394 each host species separately using network plots produced in *igraph*[73]. We calculated modularity of
395 the class-level microbial networks comprising both positive and negative interactions using the
396 modularity() function after greedy clustering implemented in the *igraph* package. We resampled 90%
397 of nodes in each network 1000 times to build distributions of modularity with which to quantify
398 differences among animal classes. We used binomial GLM to test the hypothesis that the proportion
399 of positive edges (correlations) varies by host class.

400

401 ## ACKNOWLEDGEMENTS

402 We would like to thank Miran Aprahamian and Chris Williams (Environment Agency) for providing
403 samples, as well as BEI Resources, the US Forest Service and the University of Wisconsin-Madison
404 for providing mock communities. We are grateful to Devenish Nutrition for funding the red deer
405 research at Dowth Hall, Co. Meath, Ireland. The sampling of shrews was funded by a Heredity
406 Fieldwork Grant awarded by the Genetics Society. The collection of dove and pigeon faecal samples
407 was jointly funded by the Royal Society for the Protection of Birds and Natural England through the
408 Action for Birds in England (AfBiE) partnership. Fieldwork enabling collection of avian faecal samples
409 from Lincolnshire was funded by The Royal Society Research Grant RG170086 to JCD. Small
410 mammal sampling in the Chernobyl Exclusion Zone was supported by the TREE
411 (https://tree.ceh.ac.uk/) and RED FIRE (https://www.ceh.ac.uk/redfire) projects. TREE was funded by
412 the Natural Environment Research Council (NERC), Radioactive Waste Management Ltd. and the
413 Environment Agency as part of the RATE Programme; RED FIRE was a NERC Urgency Grant.
414 Northern muriqui research at Caparaó National Park, Brazil, was funded by CAPES (BEX 1298/15-1).

415

416 ## DATA ACCESSIBILITY STATEMENT

417 Sequence data are deposited in the NCBI SRA database under BioProject numbers PRJNA593927
418 and PRJNA593220. Full analysis code has been provided as supplementary material.

419

REFERENCES

1. Archie, E. A. & Tung, J. Social behavior and the microbiome. *Curr. Opin. Behav. Sci.* **6**, 28–34 (2015).

2. McFall-Ngai, M. *et al.* Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 3229–3236 (2013).

3. Bahrndorff, S., Bahrndorff, S., Alemu, T., Alemneh, T. & Nielsen, J. L. The Microbiome of Animals : Implications for Conservation Biology. *Int. J. Genomics* **2016**, 5304028 (2016).

4. Dillon, R. J., Vennard, C. T., Buckling, A. & Charnley, A. K. Diversity of locust gut bacteria protects against pathogen invasion. *Ecol. Lett.* **8**, 1291–1298 (2005).

5. Bates, K. A. *et al.* Amphibian chytridiomycosis outbreak dynamics are linked with host skin bacterial community structure. *Nat. Commun.* **9**, 1–11 (2018).

6. Walsh, B. S., Heys, C. & Lewis, Z. Gut microbiota influences female choice and fecundity in the nuptial gift-giving species, Drosophila subobscura (Diptera: Drosophilidae). *Eur. J. Entomol.* **114**, 439–445 (2017).

7. Apprill, A. Marine animal microbiomes: Toward understanding host-microbiome interactions in a changing ocean. *Front. Mar. Sci.* **4**, 1–9 (2017).

8. Sam, Q. H., Chang, M. W. & Chai, L. Y. A. The fungal mycobiome and its interaction with gut bacteria in the host. *Int. J. Mol. Sci.* **18**, 330 (2017).

9. Enaud, R. *et al.* The Mycobiome: A Neglected Component in the Microbiota-Gut-Brain Axis. *Microorganisms* **6**, 22 (2018).

10. Frey-Klett, P. *et al.* Bacterial-Fungal Interactions: Hyphens between Agricultural, Clinical, Environmental, and Food Microbiologists. *Microbiol. Mol. Biol. Rev.* **75**, 583–609 (2011).

11. Jakuschkin, B. *et al.* Deciphering the Pathobiome: Intra- and Interkingdom Interactions Involving the Pathogen Erysiphe alphitoides. *Microb. Ecol.* **72**, 870–880 (2016).

12. van Overbeek, L. S. & Saikkonen, K. Impact of Bacterial-Fungal Interactions on the Colonization of the Endosphere. *Trends Plant Sci.* **21**, 230–242 (2016).

13. Griffiths, S. M. *et al.* Complex associations between cross-kingdom microbial endophytes and host genotype in ash dieback disease dynamics. *J. Ecol.* 1–19 (2019). doi:10.1111/1365-2745.13302

14. Ghannoum, M. A. *et al.* Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. *PLoS Pathog.* **6**, (2010).

15. Li, J. *et al.* Fecal Bacteriome and Mycobiome in Bats with Diverse Diets in South China. *Curr. Microbiol.* **75**, 1352–1361 (2018).

16. Liggenstoffer, A. S., Youssef, N. H., Couger, M. B. & Elshahed, M. S. Phylogenetic diversity and community structure of anaerobic gut fungi (phylum Neocallimastigomycota) in ruminant and non-ruminant herbivores. *ISME J.* **4**, 1225–1235 (2010).

17. Heisel, T. *et al.* High-Fat Diet Changes Fungal Microbiomes and Interkingdom Relationships in the Murine Gut. *mSphere* **2**, 1–14 (2017).

18. Wegley, L., Edwards, R., Rodriguez-Brito, B., Liu, H. & Rohwer, F. Metagenomic analysis of the microbial community associated with the coral Porites astreoides. *Environ. Microbiol.* **9**, 2707–2719 (2007).

19. Yang, S. *et al.* Metagenomic Analysis of Bacteria, Fungi, Bacteriophages, and Helminths in the Gut of Giant Pandas. *Front. Microbiol.* **9**, 1–16 (2018).

20. Kearns, P. J. *et al.* Fight fungi with fungi: Antifungal properties of the amphibian mycobiome. *Front. Microbiol.* **8**, 1–12 (2017).

21. Yeung, F. *et al.* Altered Immunity of Laboratory Mice in the Natural Environment Is Associated with Fungal Colonization. *Cell Host Microbe* 1–14 (2020). doi:10.1016/j.chom.2020.02.015

22. Lu, M., Wingfield, M. J., Gillette, N. E., Mori, S. R. & Sun, J. H. Complex interactions among host pines and fungi vectored by an invasive bark beetle. *New Phytol.* **187**, 859–866 (2010).

23. Davenport, E. R. *et al.* The human microbiome in evolution. *BMC Biol.* **15**, 1–12 (2017).

24. Knowles, S. C. L., Eccles, R. M. & Baltrūnaitė, L. Species identity dominates over environment in shaping the microbiota of small mammals. *Ecol. Lett.* **22**, 826–837 (2019).

25. Youngblut, N. D. *et al.* Host diet and evolutionary history explain different aspects of gut

473 microbiome diversity among vertebrate clades. *Nat. Commun.* **10**, 1–15 (2019).

474 26. Amato, K. R. *et al.* Evolutionary trends in host physiology outweigh dietary niche in structuring
475 primate gut microbiomes. *ISME J.* **13**, 576–587 (2019).

476 27. Song, S. J. *et al.* Comparative Analyses of Vertebrate Gut Microbiomes Reveal Convergence
477 between Birds and Bats. *MBio* **11**, 1–14 (2020).

478 28. Brooks, A. W., Kohl, K. D., Brucker, R. M., van Opstal, E. J. & Bordenstein, S. R.
479 Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host
480 Evolutionary History. *PLoS Biol.* **14**, 1–29 (2016).

481 29. Peay, K. G., Kennedy, P. G. & Bruns, T. D. Fungal Community Ecology: A Hybrid Beast with a
482 Molecular Master. *Bioscience* **58**, 799–810 (2008).

483 30. Fierer, N. Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nat.*
484 *Rev. Microbiol.* **15**, 579–590 (2017).

485 31. Rausch, P. *et al.* Comparative analysis of amplicon and metagenomic sequencing methods
486 reveals key features in the evolution of animal metaorganisms. *Microbiome* **7**, 1–19 (2019).

487 32. Stecher, B. *et al.* Like will to like: Abundances of closely related species can predict
488 susceptibility to intestinal colonization by pathogenic and commensal bacteria. *PLoS Pathog.*
489 **6**, e1000711 (2010).

490 33. Mcfrederick, Q. S., Mueller, U. G. & James, R. R. Interactions between fungi and bacteria
491 influence microbial community structure in the Megachile rotundata larval gut Interactions
492 between fungi and bacteria influence microbial community structure in the Megachile
493 rotundata larval gut. *Proc. R. Soc. B Biol. Sci.* **281**, 20132653 (2014).

494 34. Forbes, J. D., Bernstein, C. N., Tremlett, H., Van Domselaar, G. & Knox, N. C. A fungal world:
495 Could the gut mycobiome be involved in neurological disease? *Front. Microbiol.* **10**, 1–13
496 (2019).

497 35. Pareek, S. *et al.* Comparison of Japanese and Indian intestinal microbiota shows diet-
498 dependent interaction between bacteria and fungi. *npj Biofilms Microbiomes* **5**, (2019).

499 36. Hoffmann, C. *et al.* Archaea and Fungi of the Human Gut Microbiome: Correlations with Diet

and Bacterial Residents. *PLoS One* **8**, e66019 (2013).

37. Kim, W., Levy, S. B. & Foster, K. R. Rapid radiation in bacteria leads to a division of labour. *Nat. Commun.* **7**, 10508 (2016).

38. Rakoff-Nahoum, S., Foster, K. R. & Comstock, L. E. The evolution of cooperation within the gut microbiota. *Nature* **533**, 255–259 (2016).

39. Mazel, F. *et al.* Is Host Filtering the Main Driver of Phylosymbiosis across the Tree of Life? *mSystems* **3**, 1–15 (2018).

40. Muegge, B. D. *et al.* Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science (80-.).* **332**, 970–974 (2008).

41. Ingala, M. R. *et al.* Comparing microbiome sampling methods in a wild mammal: Fecal and intestinal samples record different signals of host ecology, evolution. *Front. Microbiol.* **9**, 1–13 (2018).

42. Videvall, E., Strandh, M., Engelbrecht, A., Cloete, S. & Cornwallis, C. K. Measuring the gut microbiome in birds: Comparison of faecal and cloacal sampling. *Mol. Ecol. Resour.* **18**, 424–434 (2018).

43. Leite, G. *et al.* Sa1911 – Analysis of the Small Intestinal Microbiome Reveals Marked Differences from Stool Microbiome in a Large Scale Human Cohort: Redefining the “Gut Microbiome”. *Gastroenterology* **156**, S-449-S-450 (2019).

44. Yuan, S., Cohen, D. B., Ravel, J., Abdo, Z. & Forney, L. J. Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PLoS One* **7**, (2012).

45. Weber, L., DeForce, E. & Apprill, A. Optimization of DNA extraction for advancing coral microbiota investigations. *Microbiome* **5**, 1–14 (2017).

46. Zhou, X., Nanayakkara, S., Gao, J. L., Nguyen, K. A. & Adler, C. J. Storage media and not extraction method has the biggest impact on recovery of bacteria from the oral microbiome. *Sci. Rep.* **9**, 1–10 (2019).

47. Asangba, A. E. *et al.* Variations in the microbiome due to storage preservatives are not large enough to obscure variations due to factors such as host population, host species, body site,

and captivity. *Am. J. Primatol.* **81**, 1–12 (2019).

48. Smith, D. P. & Peay, K. G. Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *PLoS One* **9**, e90234 (2014).

49. Nguyen, N. H., Smith, D., Peay, K. & Kennedy, P. Parsing ecological signal from noise in next generation amplicon sequencing. (2014).

50. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).

51. Griffiths, S. M. *et al.* Genetic variability and ontogeny predict microbiome structure in a disease-challenged montane amphibian. *ISME J.* 1 (2018). doi:10.1038/s41396-018-0167-0

52. RStudio Team. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.c (2016).

53. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/. (2017).

54. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

55. Kumar, S., Stecher, G., Suleski, M. & Hedges, S. B. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).

56. Paradis, E. & Schliep, K. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

57. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. Ggtree: an R Package for Visualization and Annotation of Phylogenetic Trees With Their Covariates and Other Associated Data. *Methods Ecol. Evol.* **8**, 28–36 (2017).

58. Bürkner, P. C. Advanced Bayesian multilevel modeling with the R package brms. *R J.* **10**, 395–411 (2018).

59. Bürkner, P. C. brms: An R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, (2017).

554 60. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis.* (Springer-Verlag New York, 2009).

555 61. Wilke, C. O. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package
556 version 1.0.0. https://CRAN.R-project.org/package=cowplot. (2019).

557 62. Kay, M. tidybayes: Tidy Data and Geoms for Bayesian Models. R package version 2.0.3,
558 http://mjskay.github.io/tidybayes/. (2020).

559 63. R Hackathon. phylobase: Base Package for Phylogenetic Structures and Comparative Data. R
560 package version 0.8.10. https://CRAN.R-project.org/package=phylobase (2020).

561 64. Keck, F., Rimet, F., Bouchez, A. & Franc, A. phylosignal: an R package to measure, test, and
562 explore the phylogenetic signal. *Ecol. Evol.* **6**, 2774–2780 (2016).

563 65. Lahti, L. & Shetty, S. Tools for microbiome analysis in R. Microbiome package version
564 1.1.10013. http://microbiome.github.com/microbiome. (2017).

565 66. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome datasets are
566 compositional: And this is not optional. *Front. Microbiol.* **8**, 1–6 (2017).

567 67. Oksanen, J. *et al.* vegan: Community Ecology Package. (2018).

568 68. McMurdie, P. J. & Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis
569 and Graphics of Microbiome Census Data. *PLoS One* **8**, e61217 (2013).

570 69. Wilman, H. *et al.* EltonTraits 1.0 : Species-level foraging attributes of the world's birds and
571 mammals. *Ecology* **95**, 2027 (2014).

572 70. Kurtz, Z. D. *et al.* Sparse and Compositionally Robust Inference of Microbial Ecological
573 Networks. *PLoS Comput. Biol.* **11**, 1–25 (2015).

574 71. Williams, R. J., Howe, A. & Hofmockel, K. S. Demonstrating microbial co-occurrence pattern
575 analyses within and between ecosystems. *Front. Microbiol.* **5**, 1–10 (2014).

576 72. Ding, B., Gentleman, R. & Carey, V. bioDist: Different distance measures. R package version
577 1.54.0. (2018).

578 73. Csárdi, G. & Nepusz, T. The igraph software package for complex network research.
579 *InterJournal Complex Syst.* **1695**, 1695 (2006).

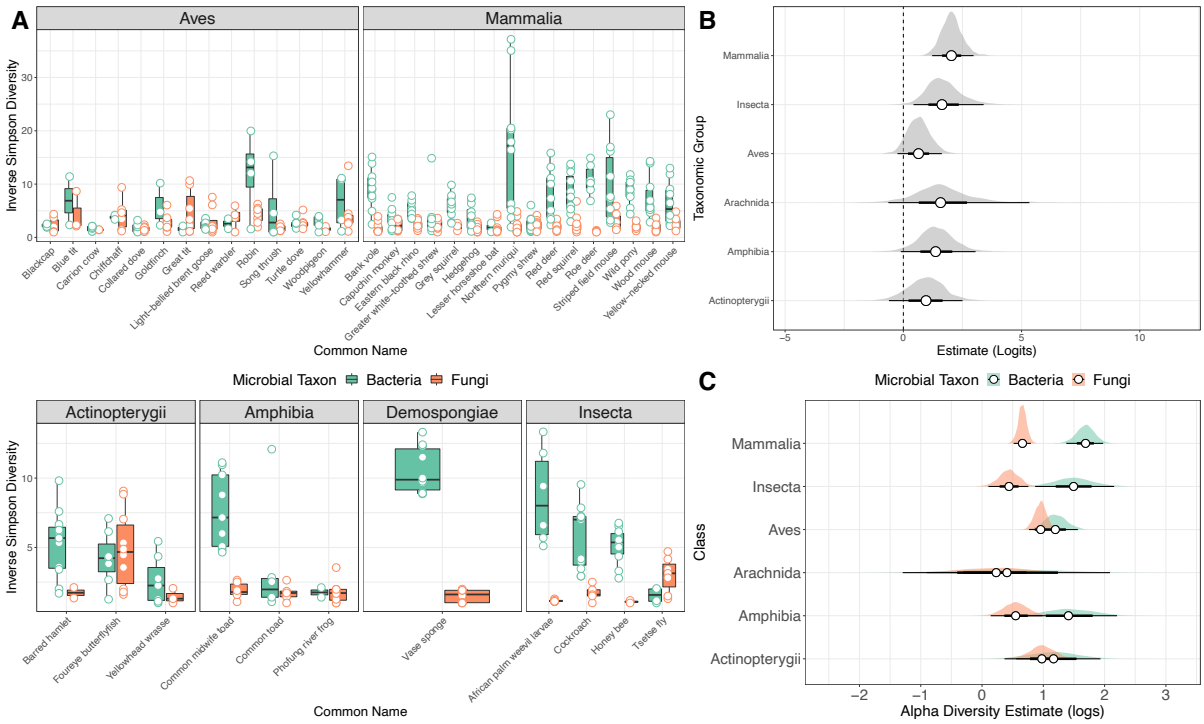


FIGURE 1

(A) Boxplots and raw data (points) of inverse Simpson indices for bacterial (green) and fungal (orange) communities across a range of host species. **(B)** Posterior estimates of a binomial mixed effects model investigating the influence of host taxonomy on the probability of a host's bacterial diversity being higher than its fungal diversity. Mammalia and Insecta had posterior estimates consistent with a >50% probability of an average animal sample having higher bacterial diversity (credible intervals do not cross zero on the link (logit) scale, equivalent to 50% on the probability scale; vertical dashed line). **(C)** Posterior estimates of a bivariate mixed effects model examining variation in average Inverse Simpson index of bacteria (green) and fungi (orange) across animal taxa. Estimates are in logs. Points are posterior means, and error bars show 66% and 95% credible intervals.

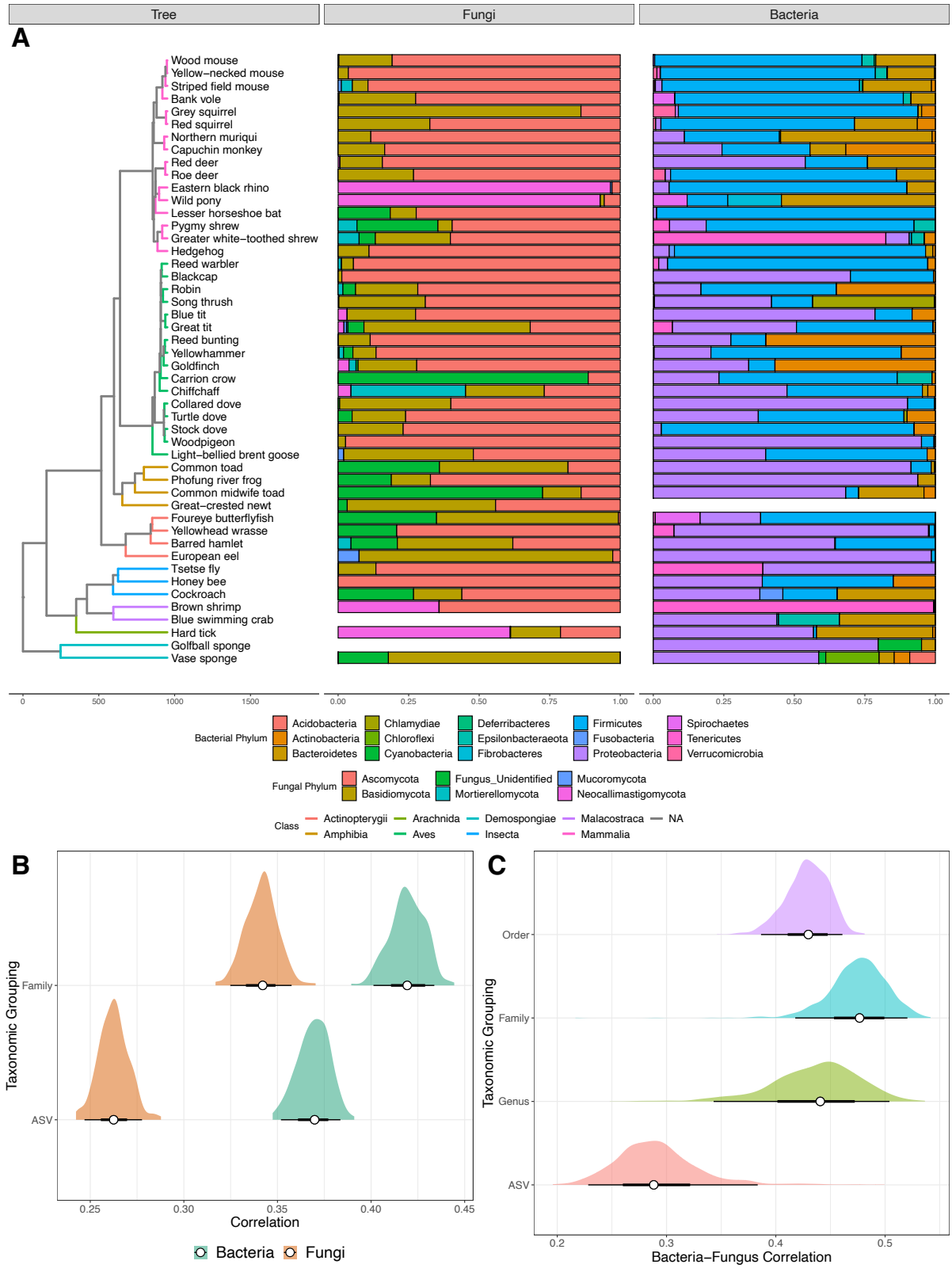
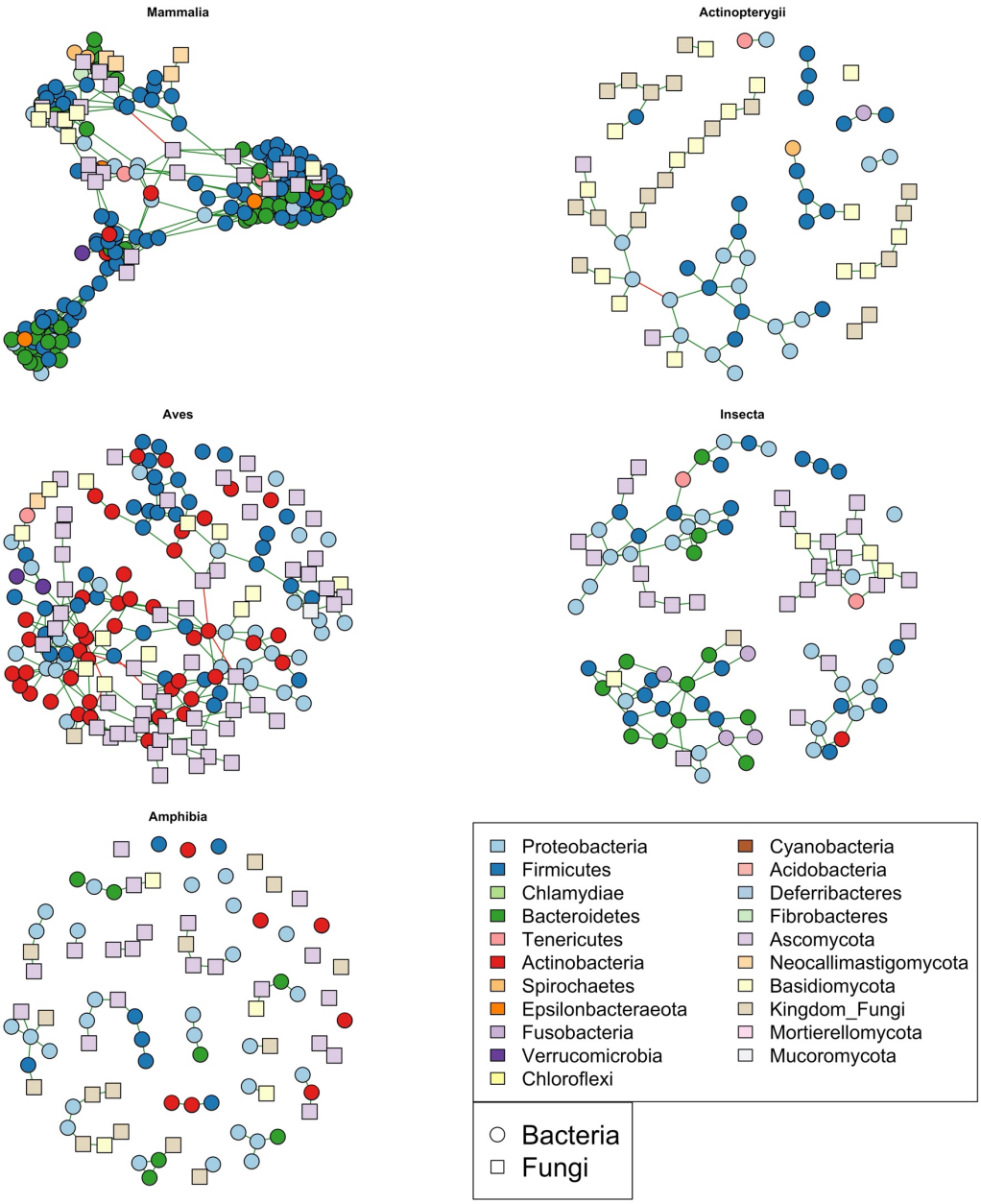


FIGURE 2

(A) Phylogenetic tree of host species, with branches coloured by class and node points coloured by order. Barplots show proportional composition of fungal and bacterial phyla for each host species,

aligned to tree tips. **(B)** Correlation of fungal and bacterial community structures (inter-sample distance) derived from Procrustes rotation on principal component ordinations of each microbial group. Microbial communities were aggregated at various taxonomic groupings (order, family, genus), or as raw Amplicon Sequence Variant (ASV) taxonomy **(C)** Correlation between matrices of inter-sample distance of microbial communities and host genetic distances generated from the phylogenetic tree in A for both bacteria (green) and fungi (orange). As for B, microbial taxonomy was either raw ASVs or grouped into family level. Aggregation to family resulted in higher correlations for both microbial groups, and the correlation was always stronger in bacteria. For both B and C, distributions of correlation values were generated using resampling of 90% of available samples for that microbial group. Empty bars mean samples were not available for a particular species and so would not have been included in the calculations in panel B.

607

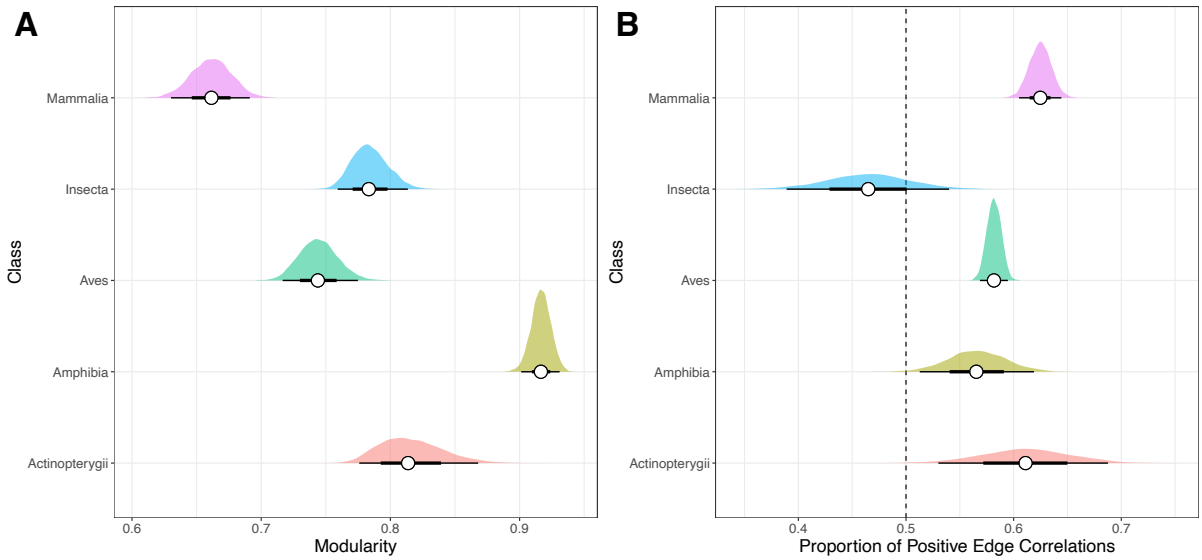


608

609 **FIGURE 3**

610 Putative microbial interaction networks between bacterial (circles) and fungal (squares) taxa, coloured
611 by microbial phylum. Networks were constructed using the R package *SpiecEasi* on CLR-transformed
612 abundance values to detect non-random co-occurrence between groups of microbes.

613



614

615 **FIGURE 4**

616 Analysis of network structural traits of five metazoan classes comprising 39 species. There were

617 significant differences in **(A)** network modularity; and **(B)** proportion of positive edges (correlations

618 between paired microbial abundance values) among classes**.** Vertical dashed line indicates equal

619 proportion of positive and negative edges.

620

TABLE 1: Details of host species and their origins, sex ratios, sample sizes and types, and storage and extraction methods for the study.

Class	Common name	Latin name	N	Sex ratio (M: F: J: unknown: N/A)	Captive or Wild	Origin	Sample Type	Collection Year	Tissue Storage	Extraction Kit
Demospongia	Vase sponge	*Ircinia campana*	10	0: 0: 0: 0: 10	Wild	Long Key, Florida, USA	Tissue (choanosome)	2014	95% ethanol	Qiagen Blood and Tissue kit with proteinase K
Demospongia	Golfball sponge	*Tethya aurantium*	10	0: 0: 0: 0: 10	Wild	Long Key, Florida, USA	Tissue (choanosome)	2014	95% ethanol	Qiagen Blood and Tissue kit with proteinase K
Arachnida	Hard tick	*Amblyomma rotundatum*	10	0: 0: 0: 10: 0	Wild	Montserrat, Caribbean	Whole organism	2014	70% ethanol	Alkaline digest and ethanol precipitation
Malacostraca	Blue swimming crab	*Portunus segnis*	5	0: 0: 0: 5: 0	Wild	Malta	Gut	2018	70% ethanol	Quigen QIAamp Fast DNA Stool Mini kit
Malacostraca	Brown shrimp	*Crangon crangon*	10	0: 0: 0: 10: 0	Wild	Liverpool, Lancashire, England	Gut	2018	Buffer AE and frozen at -20°C	Qiagen Blood and Tissue kit with proteinase K
Insecta	Cockroach	*Diploptera punctata*	11	7: 1: 0: 3: 0	Captive	Manchester Metropolitan University, Manchester, UK	Gut	2018	Liquid nitrogen and frozen at -80°C	Qiagen Blood and Tissue kit with proteinase K and lysozyme
Insecta	Honey bee	*Apis mellifera*	10	0: 10: 0: 0: 0	Wild	North West of England, UK	Gut	2016	100% ethanol and frozen at -20°C	Qiagen Blood and Tissue kit with proteinase K and lysozyme
Insecta	Tsetse fly	*Glossina fuscipes*	9	2: 7: 0: 0: 0	Wild	Patira East, Uganda	Whole organism	2019	70% ethanol	Qiagen Blood and Tissue kit with proteinase K
Insecta	African palm weevil larvae	*Rhynchophorus phoenicis*	6	0: 0: 0: 6: 0	Wild	Sapele Town, Delta State, Nigeria	Gut	2019	Frozen at -20°C	ZymoBIOMICS DNA mini kit
Actinopterygii	European eel	*Anguilla anguilla*	10	0: 0: 0: 10: 0	Wild	Cumbria, England	Gut	2009	Frozen at -20°C	Qiagen PowerSoil kit
Actinopterygii	Foureye butterflyfish	*Chaetodon capistratus*	10	0: 0: 0: 10: 0	Wild	Bocas del Toro, Bahia Almirante, Panama	Gut	2018	95% ethanol	Qiagen PowerSoil kit with proteinase K
Actinopterygii	Yellowhead wrasse	*Halichoeres garnoti*	10	0: 0: 0: 10: 0	Wild	Caye Caulker, Belize	Gut	2015	95% ethanol	Qiagen PowerSoil kit with proteinase K
Actinopterygii	Barred hamlet	*Hypoplectrus puella*	12	0: 0: 0: 12: 0	Wild	Bocas del Toro, Bahia Almirante, Panama	Gut	2018	95% ethanol	Qiagen PowerSoil kit with proteinase K
Amphibia	Common midwife toad	*Alytes obstetricans*	11	0: 0: 0: 11: 0	Captive	London Zoo, London, UK	Skin swab	2015	Frozen at -20°C	Qiagen DNEasy kit
Amphibia	Phofung river frog	*Amietia hymenopus*	10	0: 0: 10: 0: 0	Wild	Drakensberg National Park, South Africa	Tadpole mouthparts	2015	95% ethanol	Qiagen Blood and Tissue kit with proteinase K
Amphibia	Common toad	*Bufo bufo*	10	0: 0: 10: 0: 0	Wild	Norway	Whole organism	2009	70% ethanol	Phenol chlorophorm
Amphibia	Great-crested newt	*Triturus cristatus*	10	5: 5: 0: 0: 0	Wild	Lancashire, England	Toe clip	2015	70% ethanol	Phenol chlorophorm
Aves	Reed warbler	*Acrocephalus scirpaceus*	8	3: 3: 0: 2: 0	Wild	Lincolnshire, UK	Faeces	2018/19	Frozen at -20°C	Qiagen PowerSoil Pro kit
Aves	Light-bellied brent goose	*Branta bernicla hrota*	10	0: 0: 0: 10: 0	Wild	Iceland	Faeces	2017	Frozen at -20°C	Qiagen PowerSoil kit

Aves	Goldfinch	*Carduelis carduelis*	8	5: 1: 0: 2: 0	Wild	Lincolnshire, UK	Faeces	2018/19	Frozen at -20°C	Qiagen PowerSoil Pro kit
Aves	Stock dove	*Columba oenas*	10	0: 0: 0: 10: 0	Wild	East Anglia, UK	Faeces	2014	Frozen at -20°C	Qiagen QIAamp Fast DNA Stool Mini kit
Aves	Woodpigeon	*Columba palumbus*	5	0: 0: 0: 5: 0	Wild	East Anglia, UK	Faeces	2012	Frozen at -20°C	Qiagen QIAamp Fast DNA Stool Mini kit
Aves	Carrion crow	*Corvus corone*	12	0: 0: 0: 12: 0	Wild	Cumbria, UK	Gut	2019	Frozen at -20°C	Qiagen Microbiome kit
Aves	Blue tit	*Cyanistes caeruleus*	8	0: 0: 0: 8: 0	Wild	Lincolnshire, UK	Faeces	2018	Frozen at -20°C	Qiagen PowerSoil Pro kit
Aves	Yellowhammer	*Emberiza citrinella*	8	1: 1: 0: 6: 0	Wild	Lincolnshire, UK	Faeces	2018	Frozen at -20°C	Qiagen PowerSoil Pro kit
Aves	Reed bunting	*Emberiza schoeniclus*	8	4: 3: 0: 1: 0	Wild	Lincolnshire, UK	Faeces	2018	Frozen at -20°C	Qiagen PowerSoil Pro kit
Aves	Robin	*Erithacus rubecula*	8	1: 1: 0: 6: 0	Wild	Lincolnshire, UK	Faeces	2018	Frozen at -20°C	Qiagen PowerSoil Pro kit
Aves	Great tit	*Parus major*	8	3: 3: 0: 2: 0	Wild	Lincolnshire, UK	Faeces	2018/19	Frozen at -20°C	Qiagen PowerSoil Pro kit
Aves	Chiffchaff	*Phylloscopus collybita*	8	0: 1: 0: 7: 0	Wild	Lincolnshire, UK	Faeces	2018/19	Frozen at -20°C	Qiagen PowerSoil Pro kit
Aves	Collared dove	*Streptopelia decaocto*	8	0: 0: 0: 8: 0	Wild	East Anglia, UK	Faeces	2014	Frozen at -20°C	Qiagen QIAamp Fast DNA Stool Mini kit
Aves	Turtle dove	*Streptopelia turtur*	7	0: 0: 0: 7: 0	Wild	East Anglia, UK	Faeces	2014	Frozen at -20°C	Quigen QIAamp Fast DNA Stool Mini kit
Aves	Blackcap	*Sylvia atricapilla*	8	3: 2: 0: 3: 0	Wild	Lincolnshire, UK	Faeces	2018	Frozen at -20°C	Qiagen PowerSoil Pro kit
Aves	Song thrush	*Turdus philomelos*	8	0: 0: 0: 8: 0	Wild	Lincolnshire, UK	Faeces	2018	Frozen at -20°C	Qiagen PowerSoil Pro kit
Mammalia	Striped field mouse	*Apodemus agrarius*	10	5: 5: 0: 0: 0	Wild	Chernobyl Exclusion Zone, Ukraine	Faeces	2017	100% ethanol and frozen at -20°C	Invitrogen Microbiome kit
Mammalia	Yellow-necked mouse	*Apodemus flavicollis*	10	5: 5: 0: 0: 0	Wild	Chernobyl Exclusion Zone, Ukraine	Faeces	2017	100% ethanol and frozen at -20°C	Invitrogen Microbiome kit
Mammalia	Wood mouse	*Apodemus sylvaticus*	10	6: 4: 0: 0: 0	Wild	Chernobyl Exclusion Zone, Ukraine	Faeces	2017	100% ethanol and frozen at -20°C	Invitrogen Microbiome kit
Mammalia	Northern muriqui	*Brachyteles hypoxanthus*	10	0: 0: 0: 10: 0	Wild	Caparao National Park, Espirito Santo, Brazil	Faeces	2017/18	RNA Later and frozen at -20°C	Qiagen QIAamp Fast DNA Stool Mini kit
Mammalia	Roe deer	*Capreolus capreolus*	7	7: 0: 0: 0: 0	Wild	Cumbria, UK	Faeces	2019	Frozen at -20°C	Qiagen Microbiome kit
Mammalia	Red deer	*Cervus elaphus*	10	0: 0: 0: 10: 0	Wild	County Meath, Ireland	Faeces	2018	Frozen at -20°C	Qiagen QIAamp Fast DNA Stool Mini kit
Mammalia	Greater white-toothed shrew	*Crocidura russula*	10	5: 5: 0: 0: 0	Wild	Belle Ile, France	Gut	2018	100% ethanol and frozen at -20°C	Qiagen PowerSoil kit
Mammalia	Eastern black rhino	*Diceros bicornis michaeli*	10	0: 10: 0: 0: 0	Captive	Chester Zoo and Port Lympne Wild Animal Park, UK	Faeces	2011	Frozen at -20°C	Qiagen QIAamp Fast DNA Stool Mini kit
Mammalia	Wild pony	*Equus ferus caballus*	10	5: 5: 0: 0: 0	Wild	Snowdonia National Park, Wales	Faeces	2013	Frozen at -20°C	Qiagen QIAamp Fast DNA Stool Mini kit
Mammalia	Hedgehog	*Erinaceus europaeus*	12	0: 0: 0: 12: 0	Wild	Cumbria, UK	Faeces	2019	Frozen at -20°C	Qiagen Microbiome kit

Mammalia	Bank vole	*Myodes glareolus*	10	7: 3: 0: 0: 0	Wild	Chernobyl, Ukraine	Gut	2017	100% ethanol and frozen at -20°C	Invitrogen Microbiome kit
Mammalia	Lesser horseshoe bat	*Rhinolophus hipposideros*	10	3: 5: 2: 0: 0	Wild	County Kerry, Ireland	Faeces	2016	Frozen at -20°C	Zymo DNA Extraction kit
Mammalia	Capuchin monkey	*Sapajus libidinosus*	10	0: 0: 0: 10: 0	Wild	Serra Talhada, State of Pernambuco/Minas Gerais, Brazil	Faeces	2017	Frozen at -20°C	Qiagen QIAamp Fast DNA Stool Mini kit
Mammalia	Grey squirrel	*Sciurus carolinensis*	12	0: 0: 0: 12: 0	Wild	Cumbria, UK	Faeces	2019	Frozen at -20°C	Qiagen Microbiome kit
Mammalia	Red squirrel	*Sciurus vulgaris*	12	0: 0: 0: 12: 0	Wild	Cumbria, UK	Faeces	2019	Frozen at -20°C	Qiagen Microbiome kit
Mammalia	Pygmy shrew	*Sorex minutus*	10	5: 5: 0: 0: 0	Wild	Belle Ile, France	Gut	2018	100% ethanol and frozen at -20°C	Qiagen PowerSoil kit

624
625

626

627 **TABLE 2**

628 PERMANOVA results for (a) fungi and (b) bacteria of factors explaining variation in microbial
629 community structure. Terms were added in the order shown in the table to marginalise effects of
630 sample storage and preparation protocols before calculating % variance explained for taxonomic
631 groupings. Species ID was the dominant source of variation in the data for both taxonomic groups, but
632 there were also strong effects of sample storage and wet lab protocol, particularly for bacteria.

633

(a) FUNGI				Taxonomic Effects Only		
Predictor	**df**	**R2**	**p value**	**df**	**R2**	**p value**
Sample Type	7	0.05	0.001			
Tissue Storage	5	0.04	0.001			
Extraction Kit	7	0.07	0.001			
Class	2	0.02	0.001	6	0.05	0.001
Order	6	0.05	0.001	13	0.12	0.001
Species	18	0.09	0.001	26	0.14	0.001
Residuals	303	0.68		303	0.68	

(b) BACTERIA				Taxonomic Effects Only		
Predictor	**df**	**R2**	**p value**	**df**	**R2**	**p value**
Sample Type	6	0.06	0.001			
Tissue Storage	6	0.16	0.001			
Extraction Kit	7	0.12	0.001			
Class	2	0.02	0.001	6	0.09	0.001
Order	6	0.09	0.001	12	0.21	0.001
Species	18	0.12	0.001	27	0.27	0.001
Residuals	273	0.42		273	0.42	

634

635