Bird life history traits influence the diversity of their associated microbiomes

Corey T. Callaghan[1,*], Stephanie D. Jurburg[2,3]

[1] Department of Wildlife Ecology and Conservation, Fort Lauderdale Research and Education Center, University of Florida, Davie, FL 33314-7719

[2] Department of Environmental Microbiology, Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany

[3] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Germany

*corresponding author: c.callaghan@ufl.edu

Abstract

Background: The avian microbiome plays an essential role in host development, health, and behavior, but while microbiomes of captive birds have been extensively studied, little is known about how life history traits influence the resident microbial diversity of individuals and of species in wild birds. Host traits may shape their associated microbiomes by modulating the exposure of the host to microbes (e.g., through dispersal), or by selecting or removing subsets of the community, and they can affect the diversity of individuals (alpha diversity) or of the entire population (gamma diversity). To explore the relationship between interspecific traits and microbiome diversity in birds, we synthesized 773 microbiome samples and host trait data across 133 bird species and explored whether traits related to exposure to conspecifics (flock size, global abundance), environmental microbiomes (trophic level, primary habitat, primary lifestyle, body mass), or describing the range of exposure to microbes, or dispersal, (habitat breadth, range size) influence interspecific differences in individual or population-level diversity in bird-associated microbiomes.

Results: We found that traits related to exposure to environmental microbiomes (habitat, primary lifestyle, trophic level), and global abundance were the strongest predictors of differences in the composition of the bird microbiomes across species. Furthermore, we found that traits related to microbiome dispersal (range size and habitat breadth) were positively related to gamma, but not alpha diversity, highlighting that dispersal-related traits may be acting on the population level. Traits related to exposure to conspecifics were negatively related to alpha and gamma diversity, suggesting that social exposure is not a mechanism for microbial dispersal into hosts. Finally, we

found higher richness, but evidence of biotic homogenization in the microbiomes of birds inhabiting human modified systems.

Conclusions: Our study demonstrates the importance of studying interspecific differences in microbial diversity to understand the ecological drivers of host-associated microbiomes, and highlights the potential of syntheses approaches for doing so.

Keywords: Microbial diversity; trait-based ecology; 16S rRNA gene amplicon sequencing; host-microbiome relationships; synthesis

Introduction

Microbes are fundamentally important to the form and function of vertebrates (Hird 2017), working to defend them against pathogens (van der Waaij 1989; Troha and Ayres 2020), modulate behavior (Barratt et al. 2017; Trevelline and Kohl 2022), aid in digestion (Lesser and Molbak 2009; Kohl 2012), and influence nutrition (Sharpton 2018). Microbiomes are ecological communities with their own complex and dynamic interactions (Kodera et al. 2022), but they are also driven by interactions with the host and the host's environment. The microbiota of an organism can affect an individual's development, and consequently, its fitness (Rosengaus et al. 2011; Kohl et al. 2018). Conversely, the host's physiology (Rawls et al. 2006), behavior (Sarkar et al. 2020), development (Jurburg et al. 2019), and environment (Schreuder et al. 2020; Alberdi et al. 2021) can modulate the resident microbiota, creating complex feedbacks (Contijoch et al. 2019).

Understanding what shapes the microbiome of an individual can have important implications for both maintaining the stability of that microbiome and linking this to potential changes in the population of that species. Phylosymbiosis, the phylogenetic signal exhibited by host-associated microbiomes (i.e., microbiomes of more closely related species are more similar) has been repeatedly observed across animals, and attributed to selection by the host's traits, which shift as the host evolves, gradually modulating the resident microbiota (Moran and Sloan 2015; Mazel et al. 2018; Mallot and Amato 2021). This is supported by the repeated finding of convergence in the host-associated microbiome's composition across species following the convergence of their traits (Song et al. 2020). However, while host traits are responsible for this filtering, most research into the relationship between host life history traits and the gut microbiome across host

species has been limited to dietary preferences (e.g., Youngblut et al. 2019; Bodawatta et al. 2021; Mallot and Amato 2021; Levin et al. 2021; Cho and Lee 2021). Host traits can directly or indirectly shape their associated microbiomes by modulating the exposure of the host to microbes (e.g., through dispersal), or by selectively removing or selecting for subsets of the microbial community (Kohl 2020). Importantly, some life history traits like host diet may directly affect the diversity of individual animals (alpha diversity), while others, like habitat breadth may act on the microbial diversity across a population of hosts (gamma diversity), but to date, this distinction has not been tested across species.

Due to the diversity in life history strategies, global distribution, and the relative ease with which they can be sampled, wild birds represent a significant model group to further our understanding of how microbiomes are shaped (Bodawatta et al. 2022). Consequently, interest in the microbiota of birds has grown rapidly in recent years (Grond et al. 2018; Song et al. 2020; Bodawatta et al. 2022). Despite widespread declines in bird populations (e.g., Rosenberg et al. 2019), the importance of microbial diversity in maintaining populations is unknown (Grond et al. 2018). Most extant research into the host-associated microbiome has revealed substantial intraspecific variation in bird microbiomes (Song et al. 2020). In contrast, studies of the interspecific differences in microbiomes across bird hosts can reveal the role of life histories or the environment shaping the microbiome, but are rare, generally focusing on a subgroup of birds (Capunitan et al. 2020; Bodawatta et al. 2021) or including several bird species within broader studies of vertebrate-associated microbiomes (e.g., Youngblut et al. 2019). The recent surge in sequence-based studies has created a rich reservoir of microbiome data for a wide range of bird species, which can be explored within a synthesis framework.

The composition and diversity of the bird gut microbiome is likely determined by a combination of host physiology and evolutionary history, diet, and behavior (Bodawatta et al. 2022). When examined on their own, bird-associated microbiomes do not exhibit strong phylogenetic signals (Song et al. 2020; Capunitan et al. 2020; Mallott and Amato 2021), in contrast to mammals (Nishida and Ochman 2017; Clayton et al. 2018). Similarly, diet has been shown to be a key determinant of the gut microbiomes of vertebrates (Hicks et al. 2018; Youngblut et al. 2019), but there have been mixed findings for birds, with one study finding no effect of diet among birds (Song et al. 2020), and others finding diet to be a main driver of the bird gut microbiome (Hird et al. 2014; Waite and Taylor 2014). A species' preferred habitat likely constrains the environments from which colonizing microbes may be recruited. Host physiology, and especially body size, has been shown to relate to microbial diversity, with one study showing that among New Guinean passerines, larger species have more homogenous microbiomes (Bodawatta et al. 2021), and another finding a weak, negative correlation between body size and microbial richness within Passeriformes (Herder et al. 2023). Recent studies have revealed the role of social interactions in modulating the microbiomes of wild mice (Raulo et al. 2021) and chimpanzees (Moeller et al. 2016), and thus it is possible that similar life history traits in birds, such as flock size, may also modulate their microbiome.

Here, we apply a trait-based approach, using a suite of morphological and life history traits to explore the relationship between interspecific traits and microbiome diversity in birds. Specifically, we integrate and synthesize microbial data with trait data across 133 species of birds to test for interspecific differences in individual or population-level microbial diversity

among birds. We use a total of 8 different traits related to exposure to conspecifics (flock size, global abundance), environmental microbiomes (trophic level, primary habitat, primary lifestyle, body mass), or describing the range of exposure to microbes, or dispersal, (habitat breadth, range size). We hypothesize that life history traits describing the host's exposure to environmental microbiomes (trophic level, primary habitat, primary lifestyle, body mass) or to the microbiomes associated to conspecifics (flock size, global abundance), will be positively related to alpha diversity and linked to beta diversity, while traits that describe the range of environments a population encounters (e.g., range size, habitat breadth) will be positively related to gamma diversity.

Methods

Microbiome data

Raw 16S rRNA gene sequence data and associated metadata were collected from the NCBI Sequence Read Archives and Dryad (Table S1). Candidate datasets were identified from the general literature, from the Earth Microbiome Project dataset (Thompson et al. 2017), and from a list of reusable 16S rRNA gene datasets (Jurburg et al. 2020). We selected studies which (a) sequenced the V4 hypervariable region of the 16S rRNA gene between base pair positions 515 and 806; (b) sampled the bird's gastrointestinal tract, including cloacal swab, feces, intestine, and oral cavity swab. As captivity can significantly alter the resident microbiota (McKenzie et al. 2017; Hird 2017), we excluded samples from captive birds — only studies of wild (i.e., non-captive) birds were included. Details of the 7 datasets used in this study including sample types and accession numbers are available in Table S1, and a map of their geographic coverage is

shown in Figure 1. We did not account for the influence of age, sex, or season in the sampling of individual hosts, as this information was seldom reported.

Sequences were reprocessed using the dadasnake wrapper (Weissbecker et al. 2020). Our conservative approach to sequence processing was designed to maximize comparability among samples, and to focus on the dominant taxa. Primers were removed with cutadapt (Martin 2011). Only forward reads were used, and sequences were trimmed to 90 base pairs (in accordance with EMP recommendations (Thompson et al. 2017), quality-filtered, and denoised with DADA2 (Callahan et al. 2016), using standard parameters. Error-learning models were trained separately for each study. Sequences were classified using the mothur (Schloss et al. 2009) naive Bayes classifier and taxonomy was assigned using the SILVA v.138 database (Quast et al. 2013). Samples were randomly subsampled to 5,000 reads per sample. The percentage of reads preserved after filtering and removal of non-bacterial ASVs is detailed in Figure S1. Data wrangling was performed in R (3.6.3) using the phyloseq package (McMurdie and Holmes 2013). We recovered 773 microbiome samples from 133 bird species (Figure S2), representing 20,471 amplicon sequence variants (ASVs, or microbial taxa).

Life history data

We compiled data from existing sources for a total of eight life history traits, chosen based on *a priori* hypotheses and predictions as well as data availability: body mass, range size, habitat breadth, flock size, trophic level, primary habitat, primary lifestyle, and global abundance. Body mass was a continuous variable, representing the size of a species, and data were sourced from Tobias et al. 2022. Flock size was a continuous variable representing the number of individuals

of a species seen together, on average, a proxy for the gregariousness of a species (Callaghan et al. 2021). Trophic level was a categorical variable represented by either carnivore, omnivore, or herbivore, depending on diet classification and data were sourced from Pigot et al. 2020; Tobias et al. 2022. Primary habitat was a categorical variable represented by either forest, woodland, shrubland, wetland, grassland, or human modified and data were sourced from Tobias et al. 2021. Primary lifestyle was a categorical variable represented by either insessorial, generalist, terrestrial, or aerial and was sourced from Tobias et al. 2022. Global abundance represented the estimated total global abundance of a species and was sourced from Callaghan et al. 2021.

Statistical analysis

Because we had *a priori* hypotheses about individual traits, we first fit models individually for each of the eight predictor variables we tested (i.e., one response and one predictor variable). In each instance, the response variable was alpha species richness as determined from the microbiome sampling described above. We used mixed effects models where the response variable was richness, and the fixed effect (i.e., the effect of primary interest) was the predictor variable. Body mass, range size, habitat breadth, global abundance, and flock size were $\log_{10}$-transformed before modeling. Because studies within our synthesis targeted different regions of the gastrointestinal tract (Grond et al. 2018), which differ in their microbiome composition (Colston and Jackson 2016; Herder et al. 2023), we adjusted for this uncertainty in our statistical models using random effect, where the random effect for sample type (N=11) was nested within a random effect for study (N=7). We additionally used a random effect for species, as individuals of the same species were sampled more than once. Models were fit using a Poisson error

distribution in a Bayesian framework using default priors with 4 chains and 4000 iterations and a warmup of 1000 using the *brms* package in R (Bürkner et al. 2017; 2021).

To complement our single-regression model approach we fit an additional model that was a multiple regression (i.e., all eight predictor variables modeled simultaneously), following the same approach as above. Additionally, while our main results focused on species richness (Hill q=0; Chao et al. 2014), we performed exploratory analyses that used inverse Simpson diversity (q=2; Chao et al. 2014) as the response variable to confirm that the effect of traits on microbial diversity was not due to the measure of microbial diversity used. These models were fit as above, with the only difference being that Simpson's diversity was $\log_{10}$-transformed, and a gaussian error distribution was used for fitting.

To understand the relationship between microbial diversity and the host population, we performed an additional analysis that used gamma richness as the response variable. We calculated gamma richness by randomly subsampling 3 microbiome samples for each species for 1,000 iterations, calculating their cumulative richness, and averaging these richness values across iterations. Species with less than triplicate samples were excluded, and averaging reduced sample sizes, requiring that some levels of categorical variables be dropped. As such, we had fewer data points for which gamma richness models were fit (Table S2). We fit the models as above using both a single regression and multiple regression approach.

To further determine the life history traits with the strongest effect on the microbiome composition, we performed forward selection of all traits on a distance-based redundancy

analysis (RDA) of Bray-Curtis dissimilarities using the R package *vegan* (Oksanen et al. 2013) and selected the four traits with the strongest explanatory power for a variance partitioning analysis. We quantified the effect of technical choices on the microbiome by partitioning the variance in Bray-Curtis dissimilarities using sample type, study, and host species as explanatory variables.

Results

Body mass, flock size, global abundance, habitat breadth, and range size were negatively related to microbial richness in individual hosts (Figure 2), with moderate statistical support for all but flock size (see supplementary Figures S3-S7). Among trophic levels, herbivores had the lowest microbial richness, while carnivores had the highest microbial richness (Figure 2; Figure S8). Furthermore, birds inhabiting human modified habitats and wetlands had more species-rich microbiomes than those in other habitats (Figure 2; Figure S9), although we acknowledge a low sample size of species with those primary habitat types. We found minimal differences in microbial richness among birds with different lifestyles, with a slightly higher diversity in aerial species compared with terrestrial, insessorial, and generalist lifestyles (Figure 2; Figure S10).

To further examine how bird microbiomes differed at the species level, we calculated gamma richness as the cumulative microbial richness across three individuals of the same species (N=56). Body mass, flock size, and range size were negatively and significantly related with gamma microbial richness (Figure 3). These relationships were stronger (i.e., larger effect size) for gamma compared with alpha richness. In contrast, habitat breadth showed a positive relationship with gamma richness. Gamma richness was greatest in carnivores and lowest in

236 omnivores and herbivores. Compared with alpha richness, opposite patterns were found in
237 primary habitat, whereby human modified habitat had the lowest value for gamma richness
238 compared with the greatest in alpha microbial richness (Figure 3).
239
240 When looking at the same traits in a multiple regression framework (i.e., all traits modeled
241 simultaneously) we found that habitat breadth had the strongest negative relationship with
242 microbial richness, suggesting that the number of habitats a species collectively uses negatively
243 influences microbial richness of individual hosts (Figure S11). Consistent with our single
244 regression approach, we also found that a species' primary habitat modulated microbiome
245 diversity, with species inhabiting human-modified environments exhibiting the greatest
246 microbial richness. We found similar results for both species richness and inverse Simpson's
247 diversity for both our single regression approach (Figure S12) and our multiple regression
248 approach (Figure S13; Figure S14) suggesting that the observed changes in alpha richness were
249 primarily driven by the dominant community.
250
251 Host species accounted for 28.7% of the variance in community composition across all samples
252 examined, with most of this variance (15%) attributed specifically to host species, and the rest of
253 this variance jointly determined by host species, sample type (e.g., feces, cloacal swab), and
254 study (10.6%), or by host species and study (3.1%; Figure S15). Of all the traits examined,
255 forward selection revealed that primary habitat, primary lifestyle, trophic level, range size,
256 habitat breadth, and flock size collectively explained a significant ($p=0.002$), but modest
257 (adjusted $R^2=0.12$) portion of the variance in community composition. Of these, habitat had the

strongest effect on microbial community composition, explaining 3.1% of the variance in community composition on its own (Figure 4).

Discussion

Using a synthesis approach, we analyzed 773 gut microbiome samples of 133 bird species to assess the relationship between bird life history traits and microbial diversity. Broadly, we expected that traits that were associated to the exposure of individual hosts to different environments, foods, and other members of the same species would lead to greater microbial richness in individual hosts (alpha diversity), while traits that described the species-level exposure would be positively related to microbial richness across the species (gamma diversity). Our analyses found mixed support for these hypotheses and reveal that host traits can have opposite effects on alpha and gamma diversity. Ultimately, our results highlight the difficulties in predicting the avian microbiome using species-level traits, potentially underscoring the importance of environmental structure and genetic mechanisms supporting microbial diversity (Kassen and Rainey 2004).

We hypothesized that indicators of how often an individual host interacted with its peers (i.e., global abundance and flock size) would be positively related to alpha diversity, as these traits are related to the number of encounters an individual may have with conspecifics with viable microbiota (i.e., dispersal). Contrary to this expectation, we found that global abundance and flock size were negatively related to both alpha and gamma diversity, suggesting that social interactions ae not strong contributors to gut microbial diversity in birds (Sarkar et al. 2020). Other factors, such as early life or idiosyncratic exposure patterns may play a more influential

role in shaping the avian microbiome, as has been shown for captive birds (Schreuder et al. 2020).

We expected host body mass to be positively related to microbial diversity due to a greater intake of food. Contrary to our expectation, we found a negative relationship between the host's body mass for both alpha and gamma diversity. This suggests that the species-area relationship, where body size serves as a proxy for area, does not hold in microbial diversity of birds, supporting previous results by Herder et al. (2023). We also hypothesized that range size and habitat breadth, which describe the range of environments an individual or population encounters, would be positively related to alpha and gamma diversity. Instead, we found that while habitat breadth negatively related to diversity at the individual level, it related positively to the population-level microbial diversity. Indeed, habitat breadth is a metric that aggregates the habitats of all the individuals within a population, highlighting a higher host-to-host microbiome variability in species that inhabit more varied habitats. This finding suggests that as a species or population inhabits a greater variety of habitats, the microbial communities within individual hosts become less diverse, but don't adopt the same composition across hosts. One possible explanation is that species with broader habitat breadth have evolved physiological or behavioral adaptations that optimize their gut microbiome for a narrower range of environmental conditions, thereby reducing alpha diversity. In other words, although habitat breadth may be high, the individuals may rely on a specific niche within those habitats that influences the microbial diversity.

In line with our hypotheses, we found that traits related to exposure to environmental microbiomes (habitat, primary lifestyle, and trophic level) along with global abundance to be the

strongest predictors of differences in the composition of the bird microbiomes. Collectively, these four traits explained nearly 10% of the variation across microbiota, highlighting the important role of the host's exposure to microbes in modulating the kinds of microbes that inhabit the gut. Among trophic levels, we found that herbivores had the lowest microbial richness, and carnivores had the highest. At the population level, the microbiome of carnivorous species was also significantly higher than in omnivores and herbivores. This contrasts with previous studies of the relationship between host diet and the microbiome across animals. Crucially, these studies had a wide range of animal hosts, most of which were mammalian (Youngblut et al. 2019; Levin et al. 2021). By focusing exclusively on birds, this work suggests that the relationship between the host diet and its associated microbiota may vary among taxonomic classes and emphasizes the importance of exploring host-microbiome patterns within taxonomic groups.

We found higher richness in birds inhabiting human-modified systems. Notably, we found evidence of biotic homogenization in species inhabiting human-modified landscapes, where individuals had the highest microbial richness, but populations had the lowest richness. Indeed, biotic homogenization is increasingly noted as a feature of human-modified environments (Eisenhauer et al. 2023), affecting microbiomes across the world (Delgado-Baquerizo et al. 2021). While this is a novel contribution, we do acknowledge the low number of species inhabiting human-modified environments in our dataset, and further testing of this hypothesis is warranted. We also found that generalist species, characterized by their ability to use diverse resources and various habitats, showed the highest levels of gamma microbial richness. This suggests that the broader ecological niche and adaptability of generalists provide them with

increased opportunities for exposure to a wider range of microbial communities, potentially leading to greater resilience. We found minimal differences in microbial richness among birds with different lifestyles, with a slightly higher diversity in aerial species compared with terrestrial, insessorial, and generalist species. These findings suggest that the ecological strategies and niche breadth of birds play a role in shaping their associated microbiomes. However, further research is needed to understand the mechanisms underlying these associations and to explore the functional implications of these microbial diversity patterns in relation to bird lifestyles. Importantly, our study used species-level traits (i.e., one value per species, per trait). A better understanding of how traits affect the bird-associated microbiota at the individual, population, and species levels requires collecting microbiome and trait data for individuals and across bird species, and is currently not available, especially for wild birds (Bodawatta et al. 2022).

Conclusions

Over the past two decades, an increasing amount of research has explored the relationship between individual animal hosts and their gut microbiomes, focusing primarily on the role of host phylogeny (Mallott and Amato 2021) and diet (Muegge et al. 2011), revealing the important roles of each of these factors in shaping microbiomes. Less is known about how other host traits relate to the gut microbiome (Mazel et al. 2018), and whether, or to what extent, species-level traits drive microbial diversity. By studying the gut microbiota of a wide range of bird species within a synthetic framework, we found that different ecological traits shape the alpha, beta, and gamma diversity of the resident microbiota, highlighting the importance of studying drivers of microbial communities at multiple scales. The contrasting effects observed at the individual and

population levels suggest that different ecological and evolutionary processes shape microbial diversity at the individual and population scales. Further research is needed to unravel the underlying mechanisms driving these patterns and to elucidate the ecological and functional implications of these relationships in the context of host-microbe interactions. Animals exist in tight associations with their resident microbiota, and understanding how host traits shape these communities is essential to characterizing animals as holobionts (Simon et al. 2019).

Data availability

All data have been described in the methods, but parts of the data and code are available in GitHub (https://github.com/coreytcallaghan/bird_traits_microbiome) and will be permanently archived in a Zenodo repository upon acceptance. All microbiome data was previously deposited by the following studies: Koehler et al. 2016; Michel et al. 2018; Connerton et al. 2018; Capunitan et al. 2020; Berlow et al. 2021; Bodawatta et al. 2021. The corresponding accession numbers are as follows (see Table S1 for details): 10.1007/s00248-020-01569-8; 10.1098/rspb.2021.0446; 10.1111/mec.15354; 10.1186/s13071-016-1607-1; 10.1186/s40168-018-0477-5; 10.1186/s40168-018-0555-8; https://earthmicrobiome.org/.

Ethics approval and Consent to participate

Not applicable.

Consent for publication

Both authors give our consent for publication.

Availability of data and materials

All data have been described in the methods, but parts of the data and code are available in GitHub (https://github.com/coreytcallaghan/bird_traits_microbiome) and will be permanently archived in a Zenodo repository upon acceptance. All microbiome data was previously deposited by the following studies: Koehler et al. 2016; Michel et al. 2018; Connerton et al. 2018; Capunitan et al. 2020; Berlow et al. 2021; Bodawatta et al. 2021. The corresponding accession numbers are as follows (see Table S1 for details): 10.1007/s00248-020-01569-8; 10.1098/rspb.2021.0446; 10.1111/mec.15354; 10.1186/s13071-016-1607-1; 10.1186/s40168-018-0477-5; 10.1186/s40168-018-0555-8; https://earthmicrobiome.org/.

Competing interests

The authors declare no competing interests.

Funding

The authors received no specific funding for this work.

Authors' contributions

CTC and SDJ conceptualized the research, jointly performed the analyses, and jointly wrote and edited the manuscript.

Acknowledgements

We thank the Chase Synthesis group at iDiv for thoughtful discussion of this manuscript.

References

Alberdi, A., Martin Bideguren, G., & Aizpurua, O. (2021). Diversity and compositional changes in the gut microbiota of wild and captive vertebrates: A meta-analysis. *Scientific Reports*, *11*(1), 22660.

Barratt, M. J., Lebrilla, C., Shapiro, H.-Y., & Gordon, J. I. (2017). The gut microbiota, food science, and human nutrition: A timely marriage. *Cell Host & Microbe*, *22*(2), 134–141.

Berlow, M., Phillips, J. N., & Derryberry, E. P. (2021). Effects of urbanization and landscape on gut microbiomes in white-crowned sparrows. *Microbial Ecology*, *81*, 253–266.

Bodawatta, K. H., Hird, S. M., Grond, K., Poulsen, M., & Jønsson, K. A. (2022). Avian gut microbiomes taking flight. *Trends in Microbiology*, *30*(3), 268–280.

Bodawatta, K. H., Koane, B., Maiah, G., Sam, K., Poulsen, M., & Jønsson, K. A. (2021). Species-specific but not phylosymbiotic gut microbiomes of new guinean passerine birds are shaped by diet and flight-associated gut modifications. *Proceedings of the Royal Society B*, *288*(1949), 20210446.

Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*, 1–28.

Bürkner, P.-C. (2019). Bayesian item response modeling in r with brms and stan. *arXiv Preprint arXiv:1905.09501*.

Callaghan, C. T., Nakagawa, S., & Cornwell, W. K. (2021). Global abundance estimates for 9,700 bird species. *Proceedings of the National Academy of Sciences*, *118*(21), e2023170118.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from illumina amplicon data. *Nature Methods*, *13*(7), 581–583.

Capunitan, D. C., Johnson, O., Terrill, R. S., & Hird, S. M. (2020). Evolutionary signal in the gut microbiomes of 74 bird species from equatorial guinea. *Molecular Ecology*, *29*(4), 829–847.

Chao, A., Gotelli, N. J., Hsieh, T., Sander, E. L., Ma, K., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, *84*(1), 45–67.

Cho, H., & Lee, W. Y. (2020). Interspecific comparison of the fecal microbiota structure in three arctic migratory bird species. *Ecology and Evolution*, *10*(12), 5582–5594.

Clayton, J. B., Gomez, A., Amato, K., Knights, D., Travis, D. A., Blekhman, R., Knight, R., Leigh, S., Stumpf, R., Wolf, T., et al. (2018). The gut microbiome of nonhuman primates: Lessons in ecology and evolution. *American Journal of Primatology*, *80*(6), e22867.

Colston, T. J., & Jackson, C. R. (2016). Microbiome evolution along divergent branches of the vertebrate tree of life: What is known and unknown. *Molecular Ecology*, *25*(16), 3776–3800.

Connerton, P. L., Richards, P. J., Lafontaine, G. M., O'Kane, P. M., Ghaffar, N., Cummings, N. J., Smith, D. L., Fish, N. M., & Connerton, I. F. (2018). The effect of the timing of exposure to campylobacter jejuni on the gut microbiome and inflammatory responses of broiler chickens. *Microbiome, 6*, 1–17.

Contijoch, E. J., Britton, G. J., Yang, C., Mogno, I., Li, Z., Ng, R., Llewellyn, S. R., Hira, S., Johnson, C., Rabinowitz, K. M., et al. (2019). Gut microbiota density influences host physiology and is shaped by host and microbial factors. *Elife, 8*, e40553.

Delgado-Baquerizo, M., Eldridge, D. J., Liu, Y.-R., Sokoya, B., Wang, J.-T., Hu, H.-W., He, J.-Z., Bastida, F., Moreno, J. L., Bamigboye, A. R., et al. (2021). Global homogenization of the structure and function in the soil microbiome of urban greenspaces. *Science Advances, 7*(28), eabg5809.

Eisenhauer, N., Angst, G., Asato, A. E., Beugnon, R., Bönisch, E., Cesarz, S., Dietrich, P., Jurburg, S. D., Madaj, A.-M., Reuben, R. C., et al. (2023). The heterogeneity–diversity–system performance nexus. *National Science Review*, nwad109.

Grond, K., Sandercock, B. K., Jumpponen, A., & Zeglin, L. H. (2018). The avian gut microbiota: Community, physiology and function in wild birds. *Journal of Avian Biology*, *49*(11), e01788.

Herder, E. A., Skeen, H. R., Lutz, H. L., & Hird, S. M. (2023). Body size poorly predicts host-associated microbial diversity in wild birds. *Microbiology Spectrum*, e03749–22.

Hicks, A. L., Lee, K. J., Couto-Rodriguez, M., Patel, J., Sinha, R., Guo, C., Olson, S. H., Seimon, A., Seimon, T. A., Ondzie, A. U., et al. (2018). Gut microbiomes of wild great apes fluctuate seasonally in response to diet. *Nature Communications*, *9*(1), 1786.

Hird, S. M. (2017). Evolutionary biology needs wild microbiomes. *Frontiers in Microbiology*, *8*, 725.

Hird, S. M., Carstens, B. C., Cardiff, S. W., Dittmann, D. L., & Brumfield, R. T. (2014). Sampling locality is more detectable than taxonomy or ecology in the gut microbiota of the brood-parasitic brown-headed cowbird (molothrus ater). *PeerJ*, *2*, e321.

Jurburg, S. D., Brouwer, M. S., Ceccarelli, D., Goot, J. van der, Jansman, A. J., & Bossers, A. (2019). Patterns of community assembly in the developing chicken microbiome reveal rapid primary succession. *MicrobiologyOpen*, *8*(9), e00821.

Jurburg, S. D., Konzack, M., Eisenhauer, N., & Heintz-Buschart, A. (2020). The archives are half-empty: An assessment of the availability of microbial community sequencing data. *Communications Biology*, *3*(1), 474.

Kassen, R., & Rainey, P. B. (2004). The ecology and genetics of microbial diversity. *Annu. Rev. Microbiol.*, *58*, 207–231.

Kodera, S. M., Das, P., Gilbert, J. A., & Lutz, H. L. (2022). Conceptual strategies for characterizing interactions in microbial communities. *Iscience*, 103775.

Koehler, A. V., Haydon, S. R., Jex, A. R., & Gasser, R. B. (2016). Cryptosporidium and giardia taxa in faecal samples from animals in catchments supplying the city of melbourne with drinking water (2011 to 2015). *Parasites & Vectors*, *9*, 1–14.

Kohl, K. D. (2012). Diversity and function of the avian gut microbiota. *Journal of Comparative Physiology B, 182,* 591–602.

Kohl, K. D. (2020). Ecological and evolutionary mechanisms underlying patterns of phylosymbiosis in host-associated microbial communities. *Philosophical Transactions of the Royal Society B, 375*(1798), 20190251.

Kohl, K. D., Brun, A., Bordenstein, S. R., CAVIEDES-VIDAL, E., & Karasov, W. H. (2018). Gut microbes limit growth in house sparrow nestlings (passer domesticus) but not through limitations in digestive capacity. *Integrative Zoology*, *13*(2), 139–151.

Leser, T. D., & Mølbak, L. (2009). Better living through microbial action: The benefits of the mammalian gastrointestinal microbiota on the host. *Environmental Microbiology*, *11*(9), 2194–2206.

Levin, D., Raab, N., Pinto, Y., Rothschild, D., Zanir, G., Godneva, A., Mellul, N., Futorian, D., Gal, D., Leviatan, S., et al. (2021). Diversity and functional landscapes in the microbiota of animals in the wild. *Science*, *372*(6539), eabb5352.

Mallott, E. K., & Amato, K. R. (2021). Host specificity of the gut microbiome. *Nature Reviews Microbiology*, *19*(10), 639–653.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal*, *17*(1), 10–12.

Mazel, F., Davis, K. M., Loudon, A., Kwong, W. K., Groussin, M., & Parfrey, L. W. (2018). Is host filtering the main driver of phylosymbiosis across the tree of life? *Msystems*, *3*(5), e00097–18.

McKenzie, V. J., Song, S. J., Delsuc, F., Prest, T. L., Oliverio, A. M., Korpita, T. M., Alexiev, A., Amato, K. R., Metcalf, J. L., Kowalewski, M., et al. (2017). The effects of captivity on the mammalian gut microbiome. *Integrative and Comparative Biology*, *57*(4), 690–704.

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data. *PloS One*, *8*(4), e61217.

Moeller, A. H., Foerster, S., Wilson, M. L., Pusey, A. E., Hahn, B. H., & Ochman, H. (2016). Social behavior shapes the chimpanzee pan-microbiome. *Science Advances*, *2*(1), e1500997.

Moran, N. A., & Sloan, D. B. (2015). The hologenome concept: Helpful or hollow? *PLoS Biology*, *13*(12), e1002311.

Muegge, B. D., Kuczynski, J., Knights, D., Clemente, J. C., González, A., Fontana, L., Henrissat, B., Knight, R., & Gordon, J. I. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science*, *332*(6032), 970–974.

Nishida, A. H., & Ochman, H. (2018). Rates of gut microbiome divergence in mammals. *Molecular Ecology*, *27*(8), 1884–1897.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R., Simpson, G. L., Solymos, P., Stevens, M. H. H., Wagner, H., et al. (2013). Package "vegan." *Community Ecology Package, Version*, *2*(9), 1–295.

Pigot, A. L., Sheard, C., Miller, E. T., Bregman, T. P., Freeman, B. G., Roll, U., Seddon, N., Trisos, C. H., Weeks, B. C., & Tobias, J. A. (2020). Macroevolutionary convergence connects morphological form to ecological function in birds. *Nature Ecology & Evolution*, *4*(2), 230–239.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596.

Raulo, A., Allen, B. E., Troitsky, T., Husby, A., Firth, J. A., Coulson, T., & Knowles, S. C. (2021). Social networks strongly predict the gut microbiota of wild mice. *The ISME Journal*, *15*(9), 2601–2613.

Rawls, J. F., Mahowald, M. A., Ley, R. E., & Gordon, J. I. (2006). Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell*, *127*(2), 423–433.

Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., Parr, M., et al. (2019). Decline of the north american avifauna. *Science*, *366*(6461), 120–124.

Rosengaus, R. B., Zecher, C. N., Schultheis, K. F., Brucker, R. M., & Bordenstein, S. R. (2011). Disruption of the termite gut microbiota and its prolonged consequences for fitness. *Applied and Environmental Microbiology*, *77*(13), 4303–4312.

Sarkar, A., Harty, S., Johnson, K. V.-A., Moeller, A. H., Archie, E. A., Schell, L. D., Carmody, R. N., Clutton-Brock, T. H., Dunbar, R. I., & Burnet, P. W. (2020). Microbial

transmission in animal social networks and the social microbiome. *Nature Ecology & Evolution*, *4*(8), 1020–1035.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., et al. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, *75*(23), 7537–7541.

Schreuder, J., Velkers, F. C., Bouwstra, R. J., Beerens, N., Stegeman, J. A., Boer, W. F. de, Hooft, P. van, Elbers, A. R., Bossers, A., & Jurburg, S. D. (2020). An observational field study of the cloacal microbiota in adult laying hens with and without access to an outdoor range. *Animal Microbiome*, *2*(1), 1–11.

Sharpton, T. J. (2018). Role of the gut microbiome in vertebrate evolution. *Msystems*, *3*(2), e00174–17.

Simon, J.-C., Marchesi, J. R., Mougel, C., & Selosse, M.-A. (2019). Host-microbiota interactions: From holobiont theory to analysis. *Microbiome*, *7*(1), 1–5.

Song, S. J., Sanders, J. G., Delsuc, F., Metcalf, J., Amato, K., Taylor, M. W., Mazel, F., Lutz, H. L., Winker, K., Graves, G. R., et al. (2020). Comparative analyses of vertebrate gut microbiomes reveal convergence between birds and bats. *MBio*, *11*(1), e02901–19.

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., et al. (2017). A communal catalogue reveals earth's multiscale microbial diversity. *Nature*, *551*(7681), 457–463.

Tobias, J. A., Sheard, C., Pigot, A. L., Devenish, A. J., Yang, J., Sayol, F., Neate-Clegg, M. H., Alioravainen, N., Weeks, T. L., Barber, R. A., et al. (2022). AVONET: Morphological, ecological and geographical data for all birds. *Ecology Letters*, *25*(3), 581–597.

Trevelline, B. K., & Kohl, K. D. (2022). The gut microbiome influences host diet selection behavior. *Proceedings of the National Academy of Sciences*, *119*(17), e2117537119.

Troha, K., & Ayres, J. S. (2020). Metabolic adaptations to infections at the organismal level. *Trends in Immunology*, *41*(2), 113–125.

Waaij, V. (1989). The ecology of the human intestine and its consequences for overgrowth by pathogens such as clostridium difficile. *Annual Review of Microbiology*, *43*(1), 69–87.

Waite, D. W., & Taylor, M. W. (2014). Characterizing the avian gut microbiota: Membership, driving influences, and potential function. *Frontiers in Microbiology*, *5*, 223.

Weißbecker, C., Schnabel, B., & Heintz-Buschart, A. (2020). Dadasnake, a snakemake implementation of DADA2 to process amplicon sequencing data for microbial ecology. *GigaScience*, *9*(12), giaa135.

Youngblut, N. D., Reischer, G. H., Walters, W., Schuster, N., Walzer, C., Stalder, G., Ley, R. E., & Farnleitner, A. H. (2019). Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. *Nature Communications*, *10*(1), 1–15.

Figures

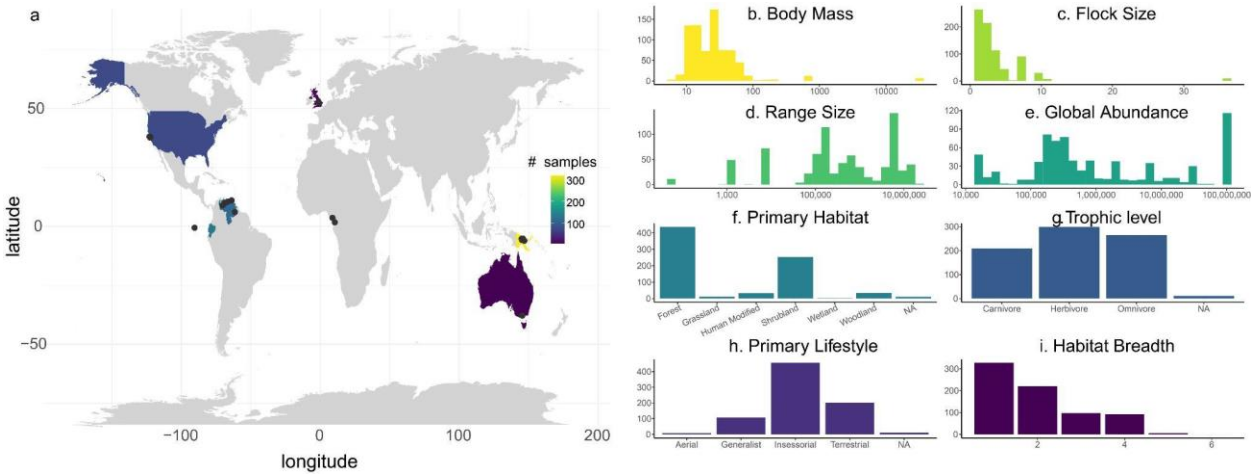

Figure 1. (a) A description of the bird samples and traits included in this study. Birds were sampled from globally distributed locations. Samples were collected in Australia, Ecuador, Equatorial Guinea, Papua New Guinea, UK, USA, and Venezuela. Countries are colored according to the number of samples collected in each location. Purple dots indicate sampling sites. Traits analyzed included a range of body, flock, and range sizes (b-d), global abundances (e), and habitat, lifestyle, and nutritional preferences (g-i).

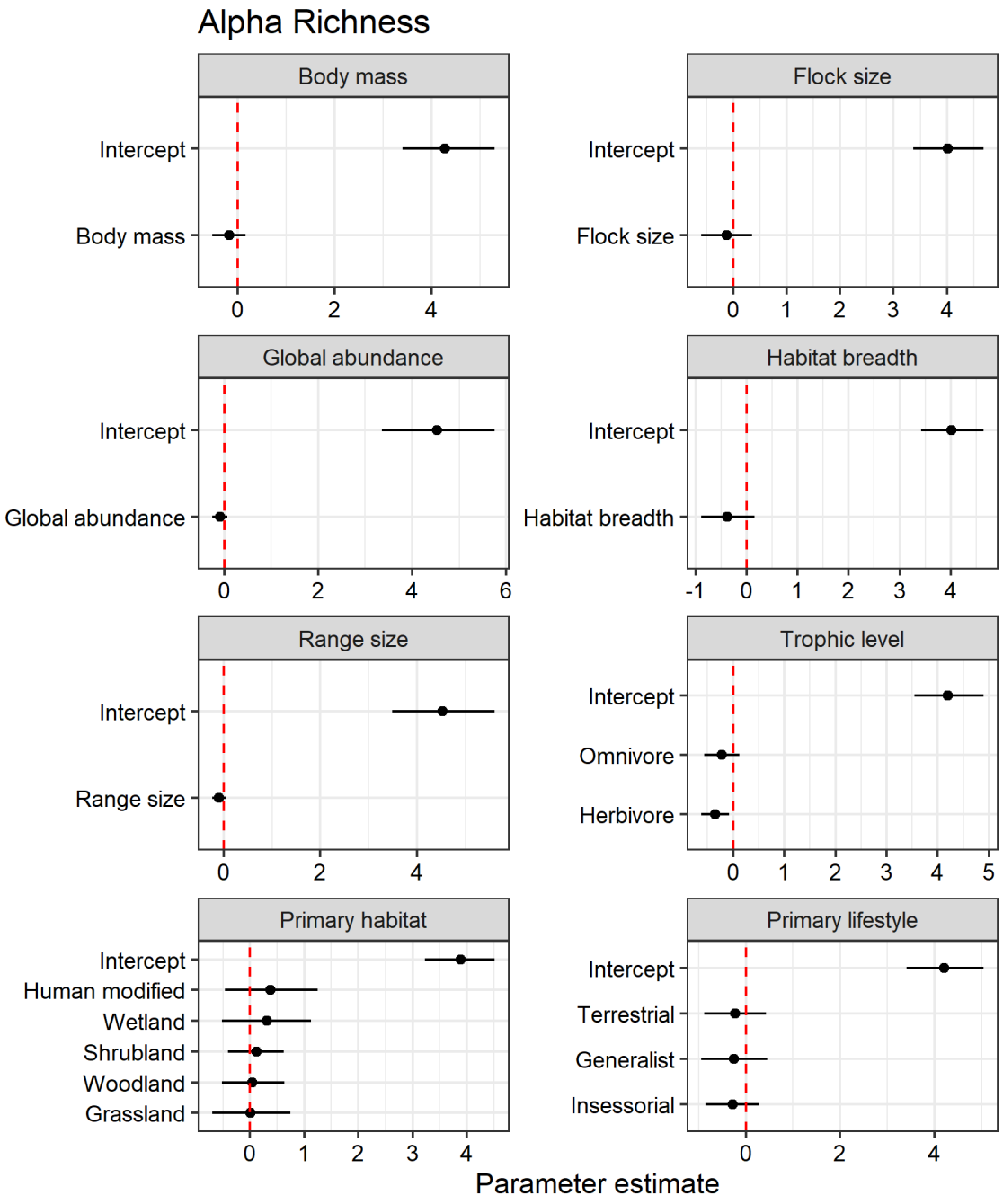

Figure 2. The results of individual regression models (N=8) where the response variable was alpha microbial species richness. Each panel represents a separate model, and the red vertical line is at zero, representing no influence of the predictor variable on the response variable. The black lines represent the 95% credible interval.

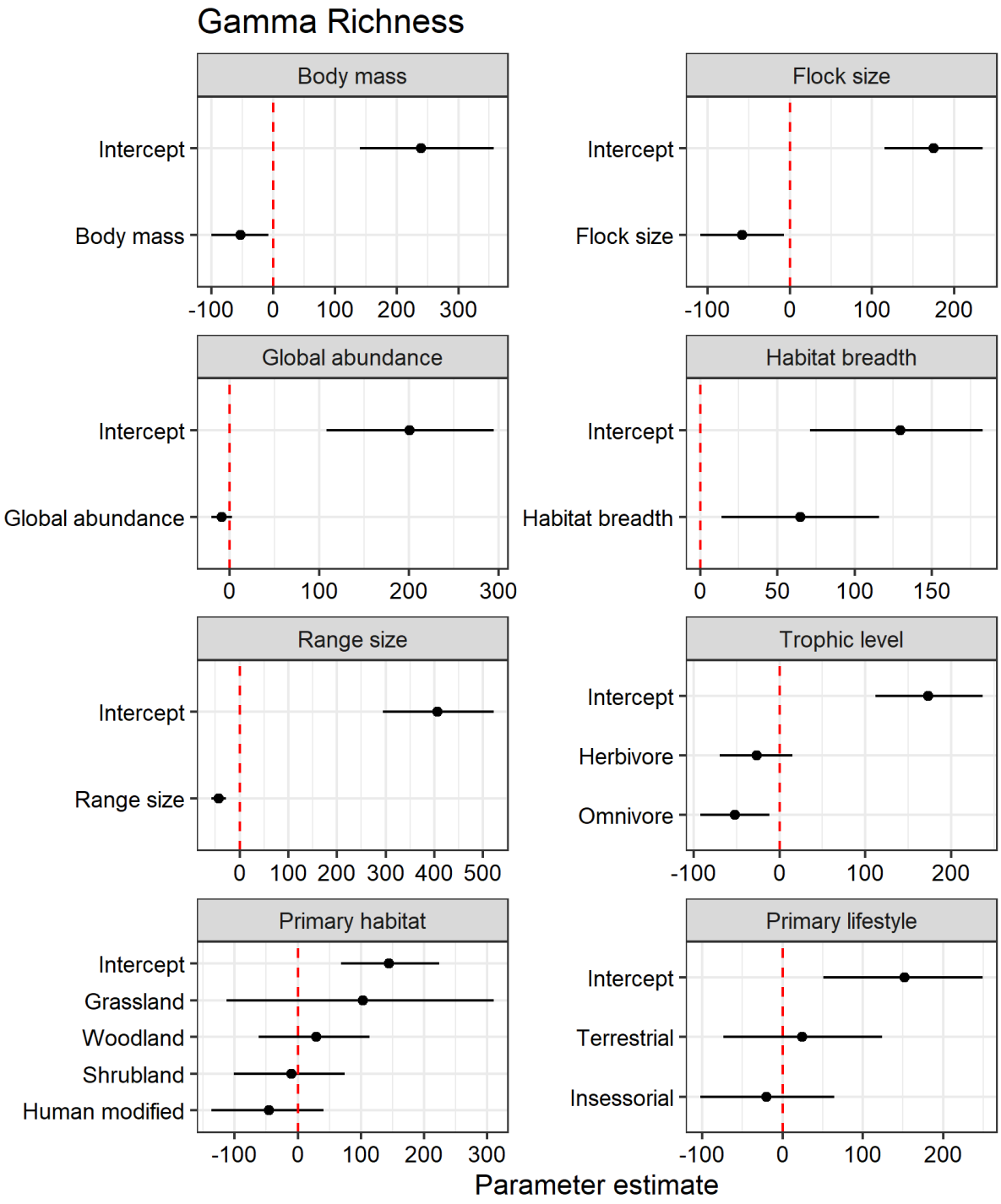

Figure 3. The results of individual regression models (N=8) where the response variable was gamma microbial species richness. Each panel represents a separate model, and the red vertical line is at zero, representing no influence of the predictor variable on the response variable. The black lines represent the 95% credible interval.

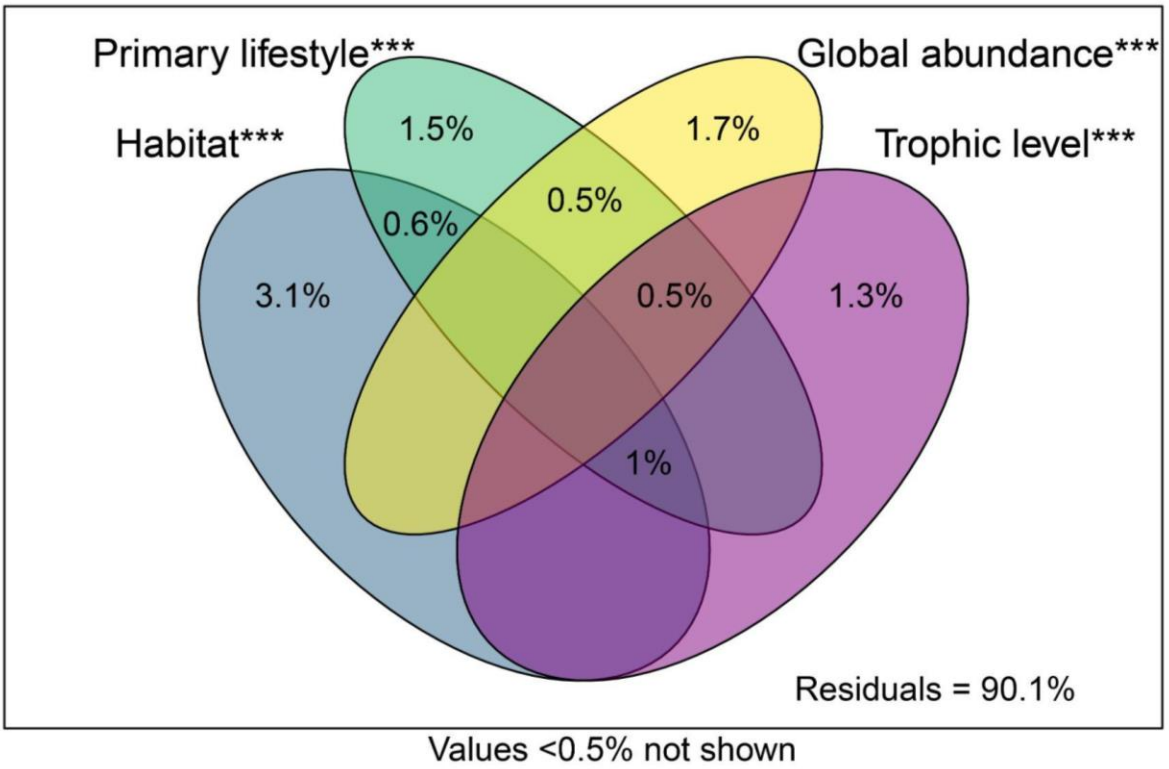

Figure 4. Variance partitioning analysis of Bray-Curtis dissimilarities. All samples which had complete trait data (n=717 samples) were included. The traits with the strongest explanatory power for the microbiome were selected using permutation-based backward and forward selection. *** p-values for permutation-based tests for the individual significance of each trait is < 0.001.

Supplementary Tables

Table S1. Metadata information for the seven studies used in the synthesis.

Study DOI	# samples	# species	mean observations per sample (pre-processing)
10.1007/s00248-020-01569-8	77	1	170989
10.1098/rspb.2021.0446	308	45	42448
10.1111/mec.15354	125	67	62611
10.1186/s13071-016-1607-1	7	1	45490
10.1186/s40168-018-0477-5	12	1	110855
10.1186/s40168-018-0555-8	131	9	38418
Earth Microbiome Project	113	9	60941

Table S2. The number of species and total observations in each respective model for single regression models (i.e., where each predictor trait was treated individually), for alpha and gamma microbial richness. For multiple regression (i.e., where all predictors were modeled simultaneously) there were 122 unique species and 717 observations from 7 studies.

Predictor	**Number of species**		**Number of observations**		**Number of studies**
	Alpha	**Gamma**	**Alpha**	**Gamma**	
Body mass	133	56	773	249	7
Range size	132	55	763	248	7
Habitat breadth	126	54	743	247	7
Global abundance	130	54	757	247	7
Flock size	132	56	772	249	7
Primary habitat	133	56	773	249	7
Trophic level	133	56	773	249	7
Primary lifestyle	133	56	773	249	7

Supplementary Figures

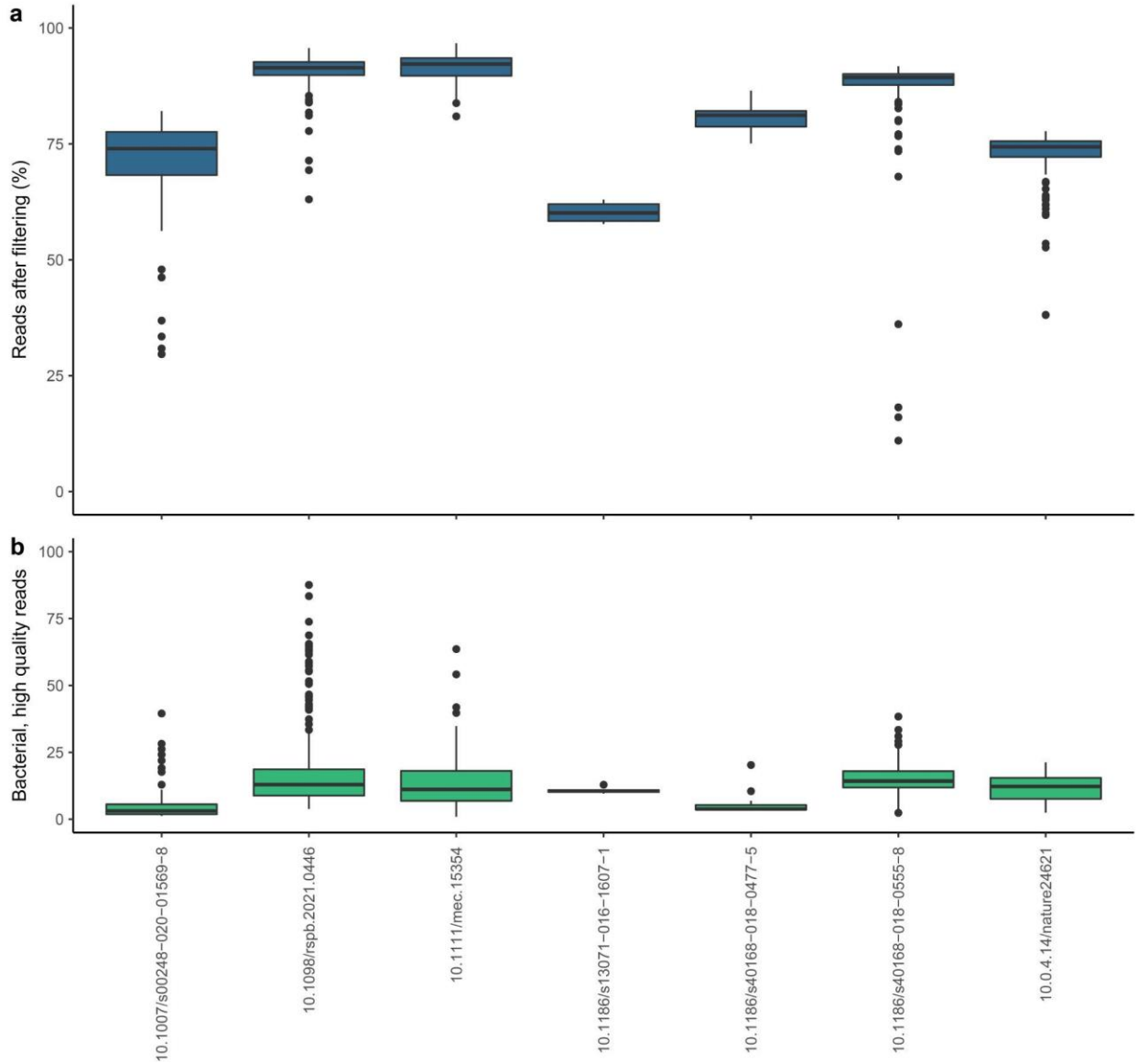

Figure S1. Proportion of reads preserved after quality filtering (a), the proportion of reads included in the synthesis study (b) for each dataset. Additional information for each study is found in Table S1.

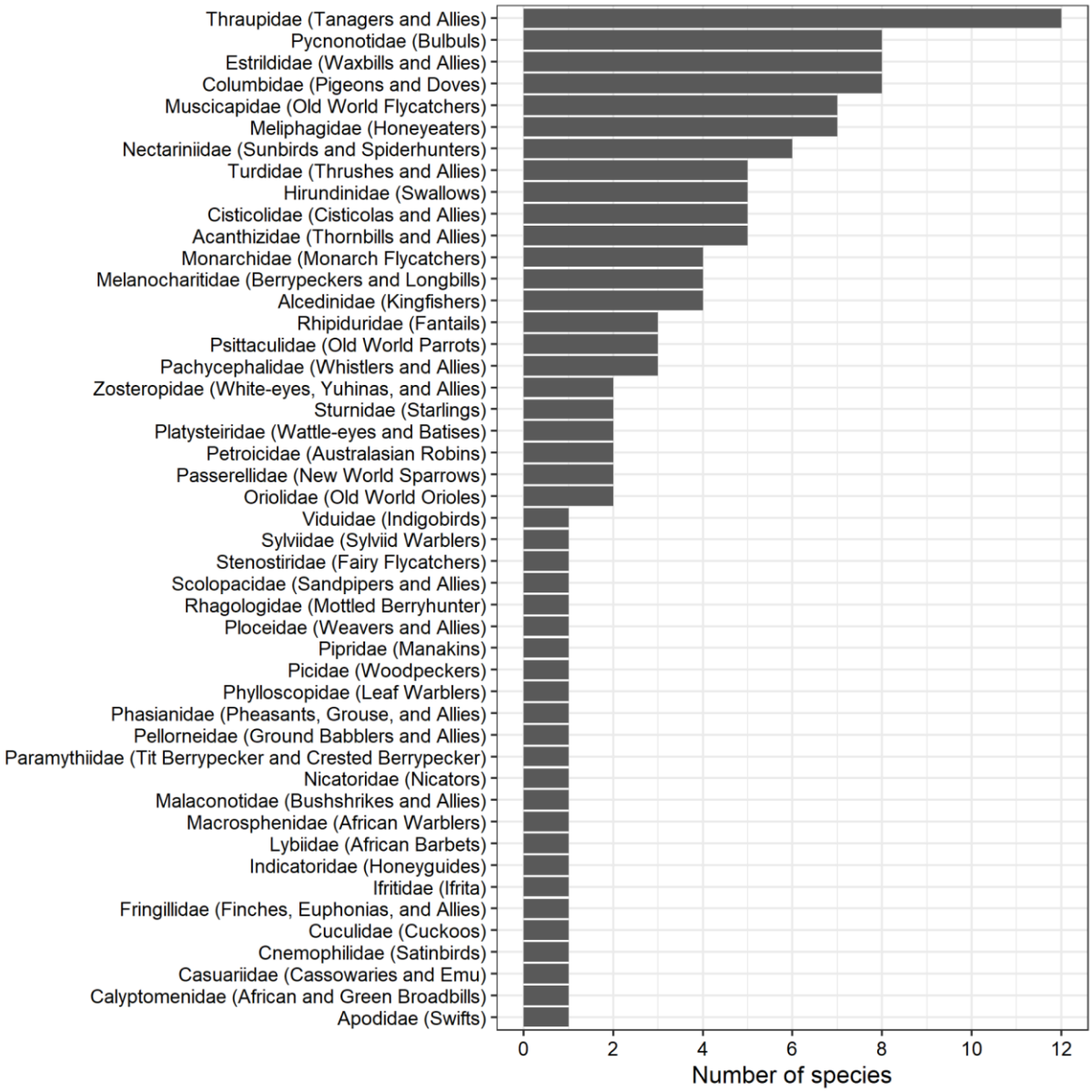

Figure S2. We had a total of 133 species included in our alpha analysis across 47 families.

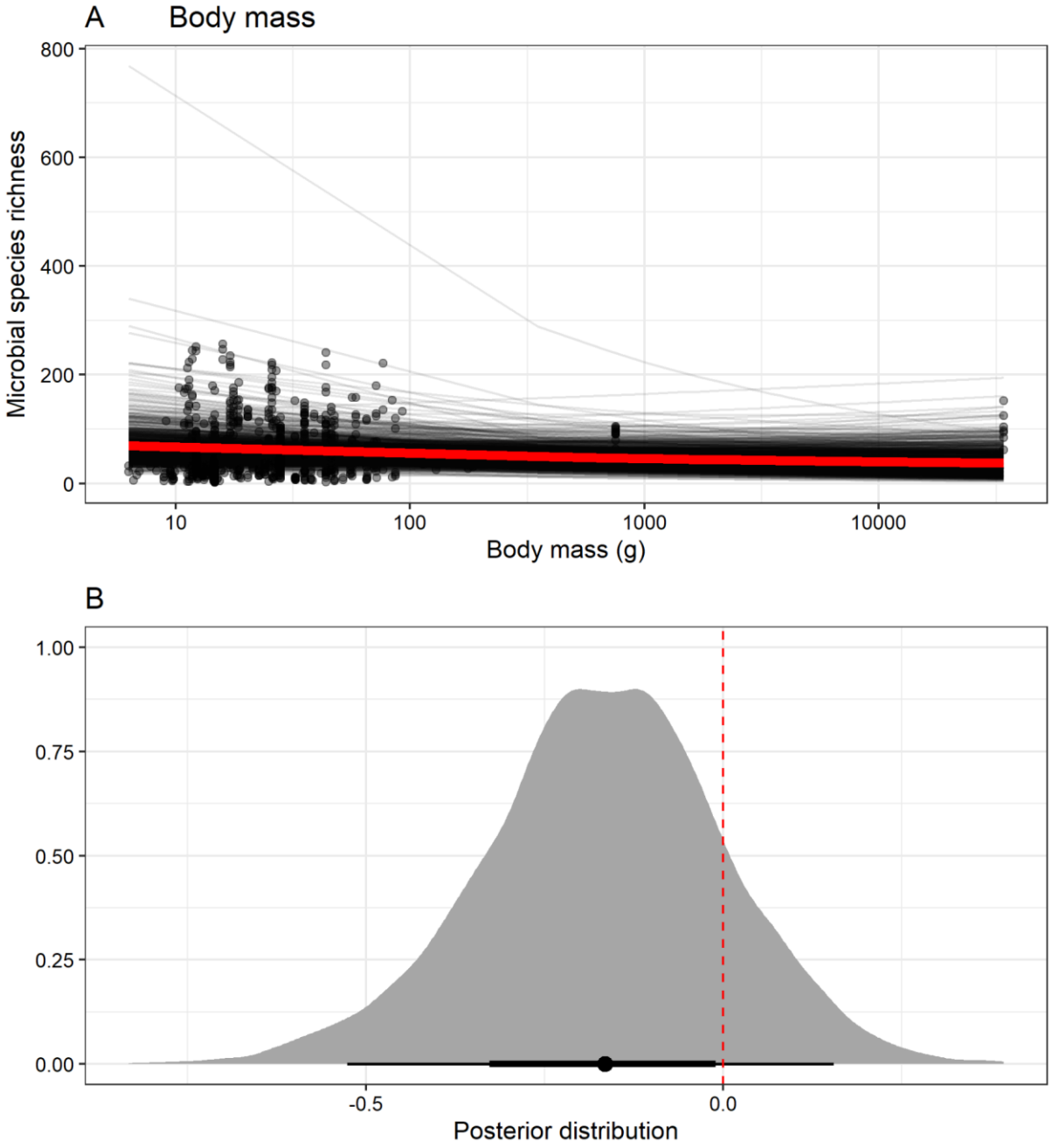

Figure S3. The results of our single regression model where alpha species richness was the response variable and body mass was the predictor variable, log10-transformed to meet modelling assumptions. Panel A represents predicted draws from our bayesian model where the red line is the average relationship and each individual line represents a separate posterior prediction, and the points represent the raw data. Panel B shows the posterior distribution of the relationship for body mass.

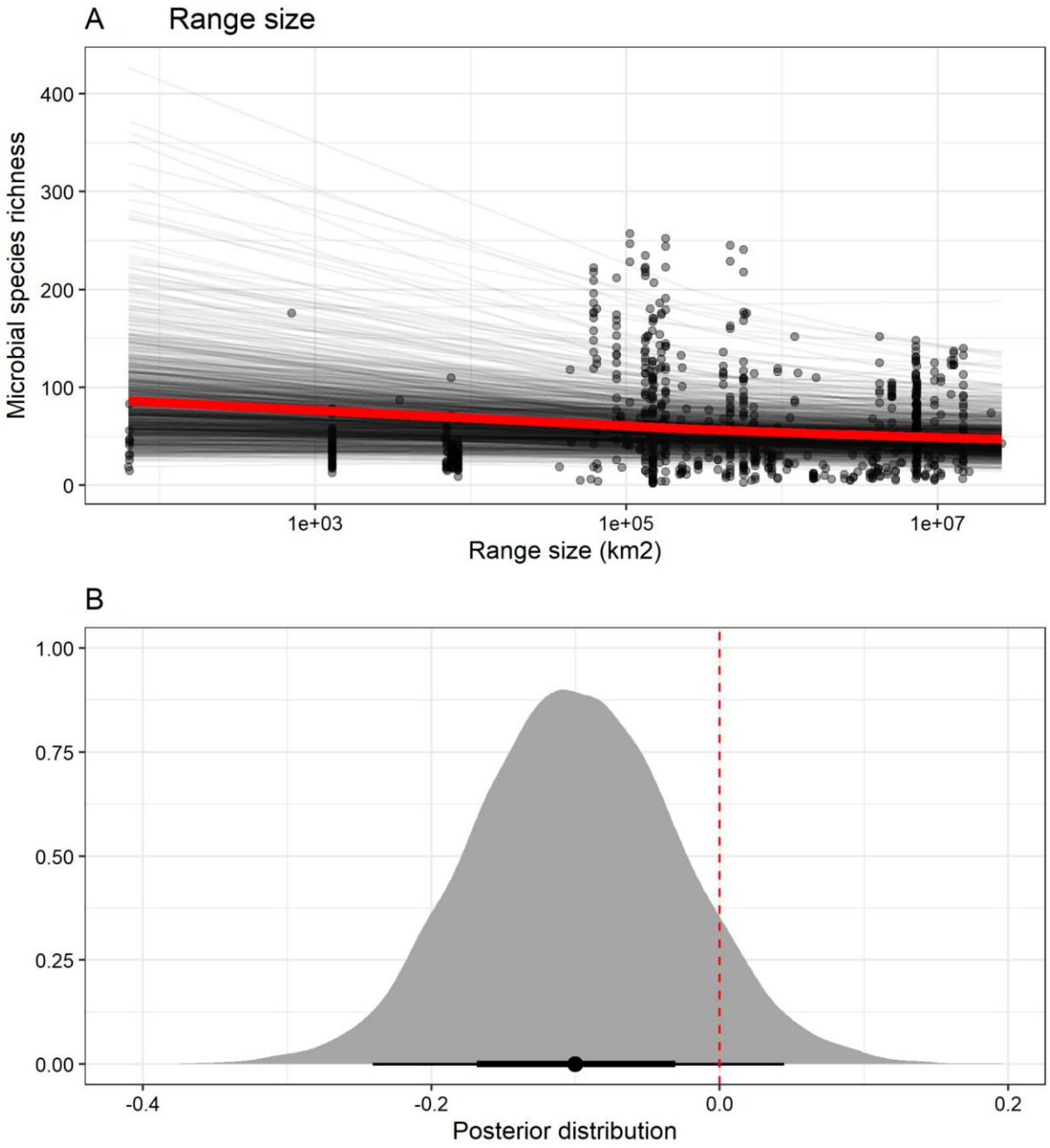

Figure S4. The results of our single regression model where alpha species richness was the response variable and body mass was the predictor variable, log10-transformed to meet modelling assumptions. Panel A represents predicted draws from our bayesian model where the red line is the average relationship and each individual line represents a separate posterior prediction, and the points represent the raw data. Panel B shows the posterior distribution of the relationship for body mass.

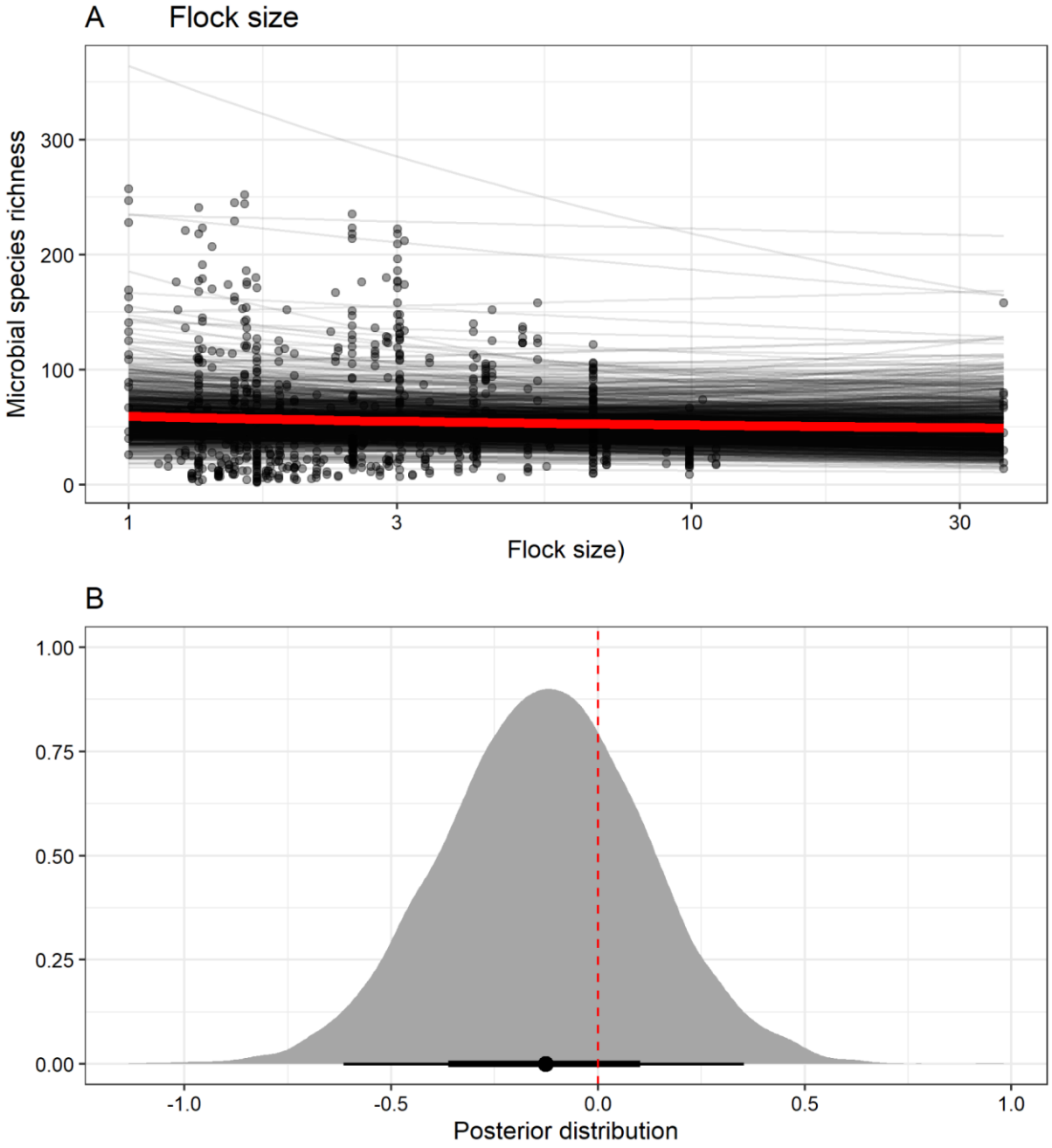

Figure S5. The results of our single regression model where alpha species richness was the response variable and body mass was the predictor variable, log10-transformed to meet modelling assumptions. Panel A represents predicted draws from our bayesian model where the red line is the average relationship and each individual line represents a separate posterior prediction, and the points represent the raw data. Panel B shows the posterior distribution of the relationship for body mass.

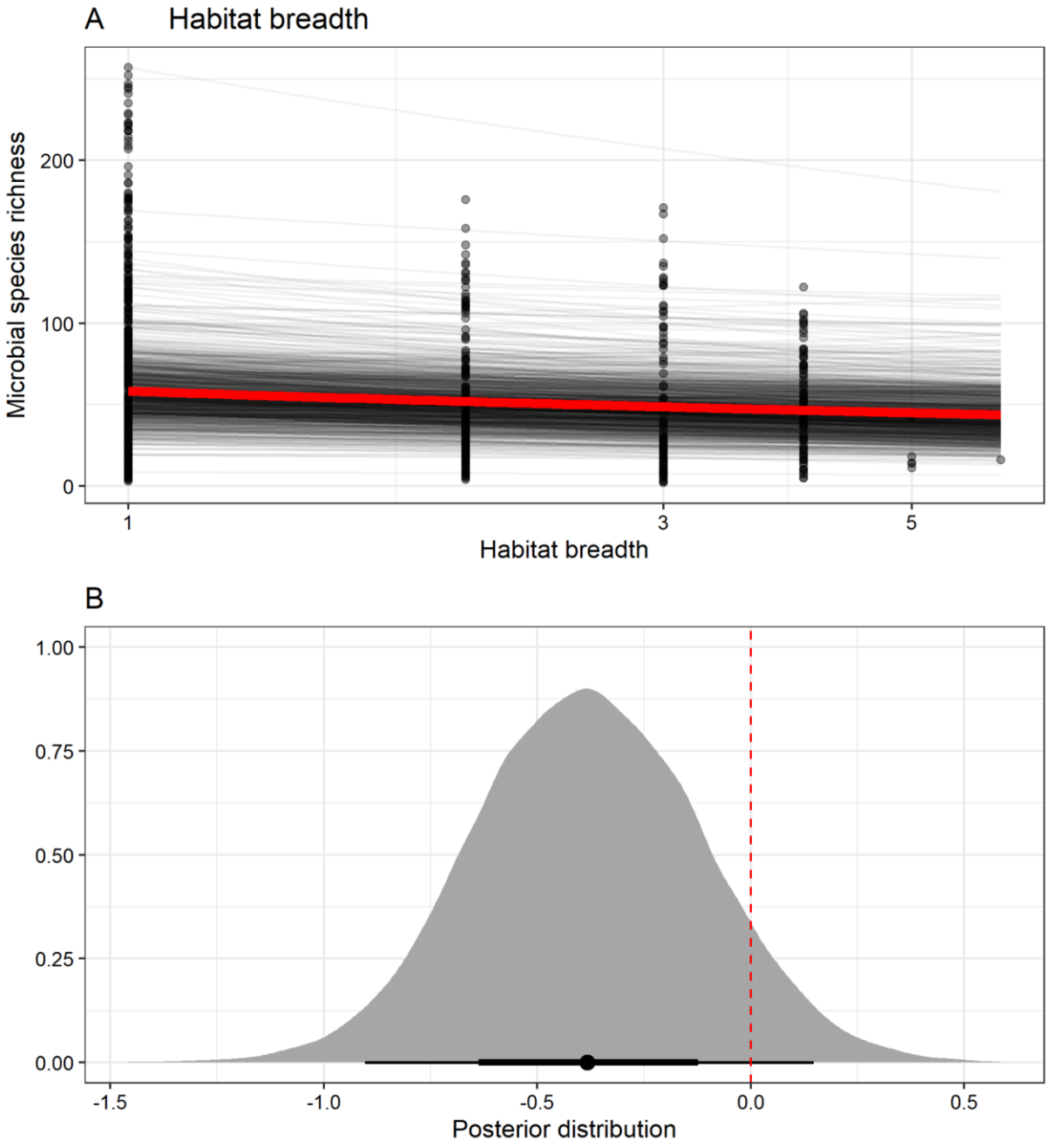

Figure S6. The results of our single regression model where alpha species richness was the response variable and body mass was the predictor variable, log10-transformed to meet modelling assumptions. Panel A represents predicted draws from our bayesian model where the red line is the average relationship and each individual line represents a separate posterior prediction, and the points represent the raw data. Panel B shows the posterior distribution of the relationship for body mass.

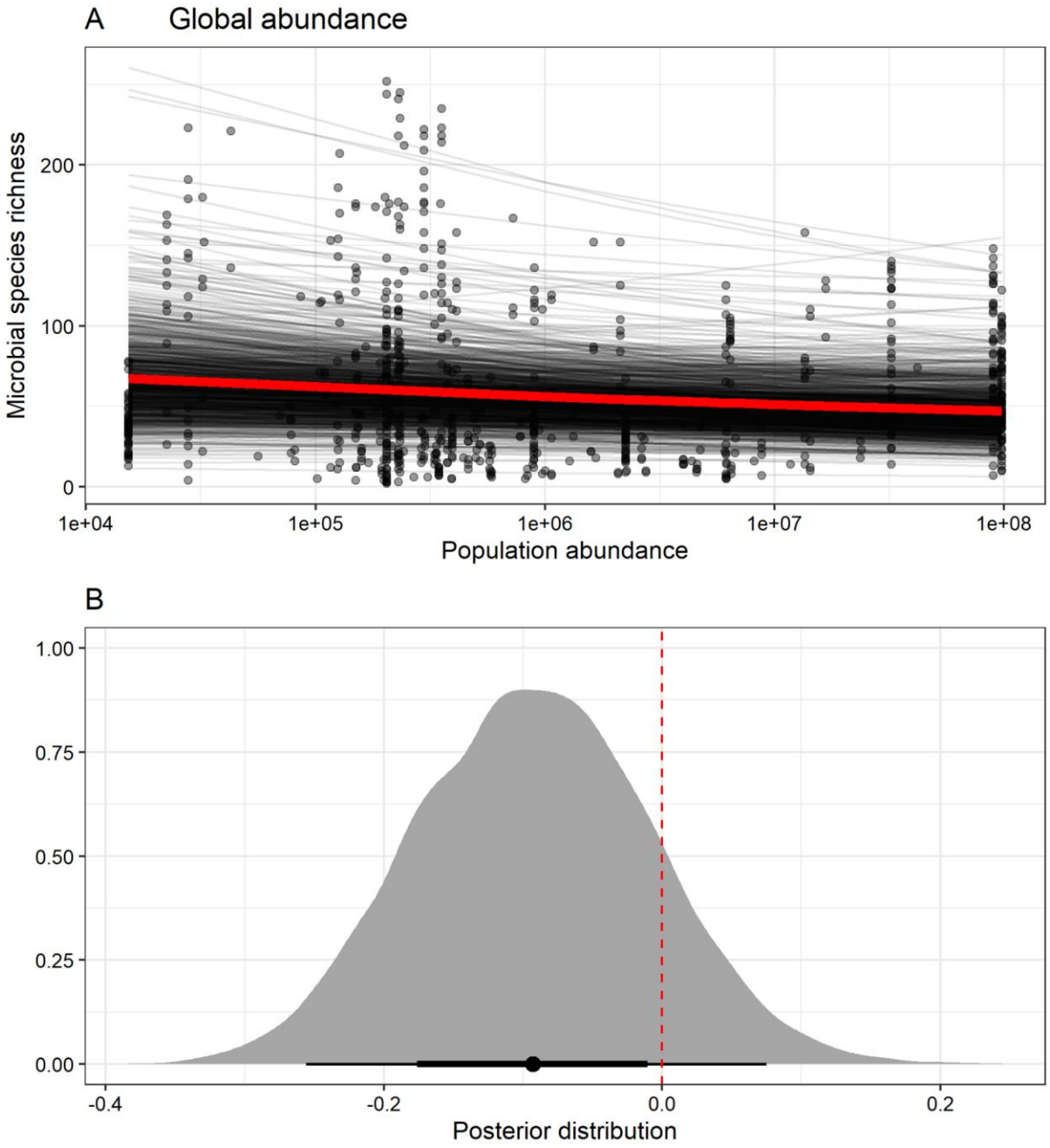

Figure S7. The results of our single regression model where alpha species richness was the response variable and body mass was the predictor variable, log10-transformed to meet modelling assumptions. Panel A represents predicted draws from our bayesian model where the red line is the average relationship and each individual line represents a separate posterior prediction, and the points represent the raw data. Panel B shows the posterior distribution of the relationship for body mass.

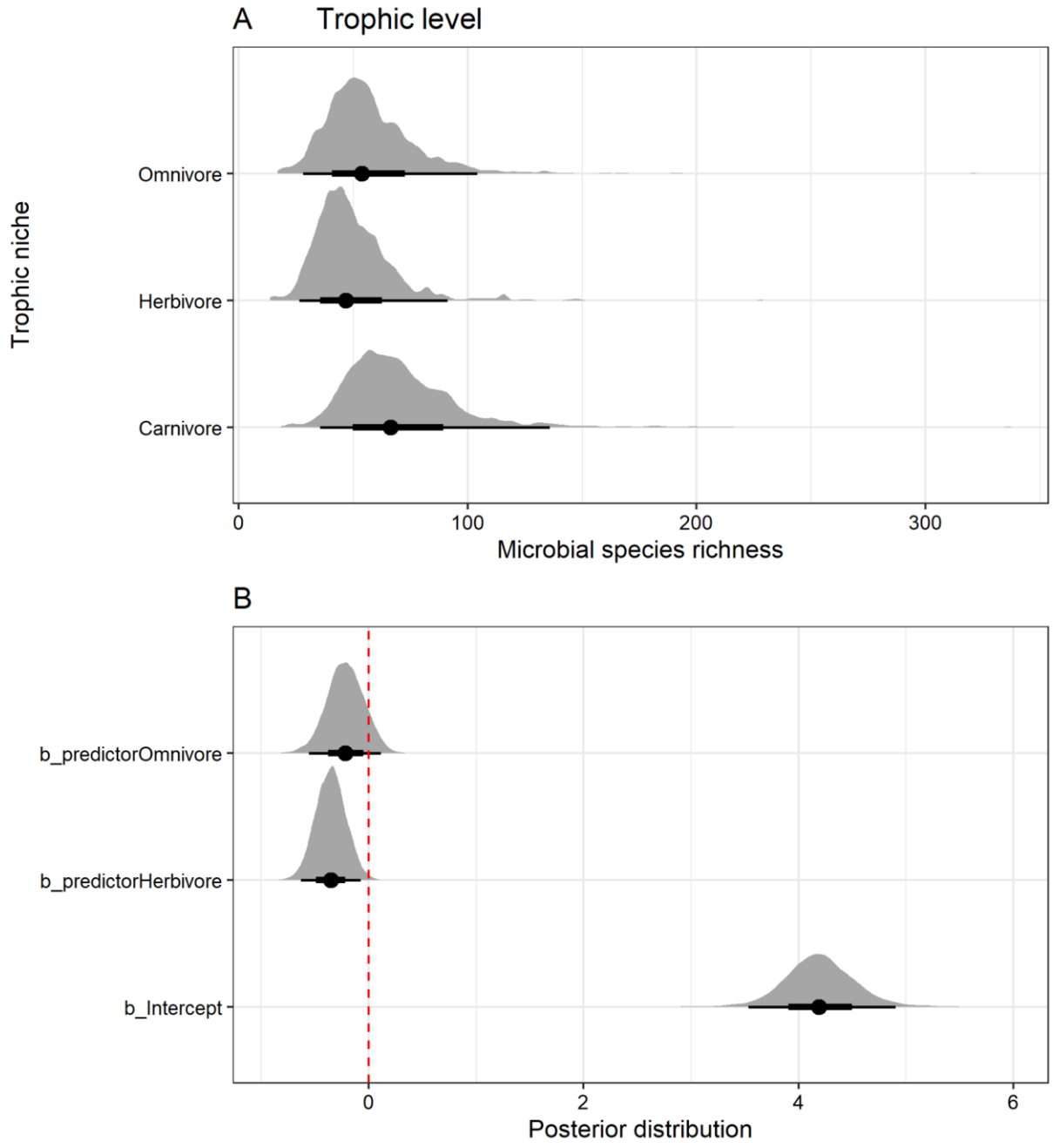

Figure S8. The results of our single regression model where alpha species richness was the response variable and body mass was the predictor variable, log10-transformed to meet modelling assumptions. Panel A represents predicted draws from our bayesian model where the red line is the average relationship and each individual line represents a separate posterior prediction, and the points represent the raw data. Panel B shows the posterior distribution of the relationship for body mass.

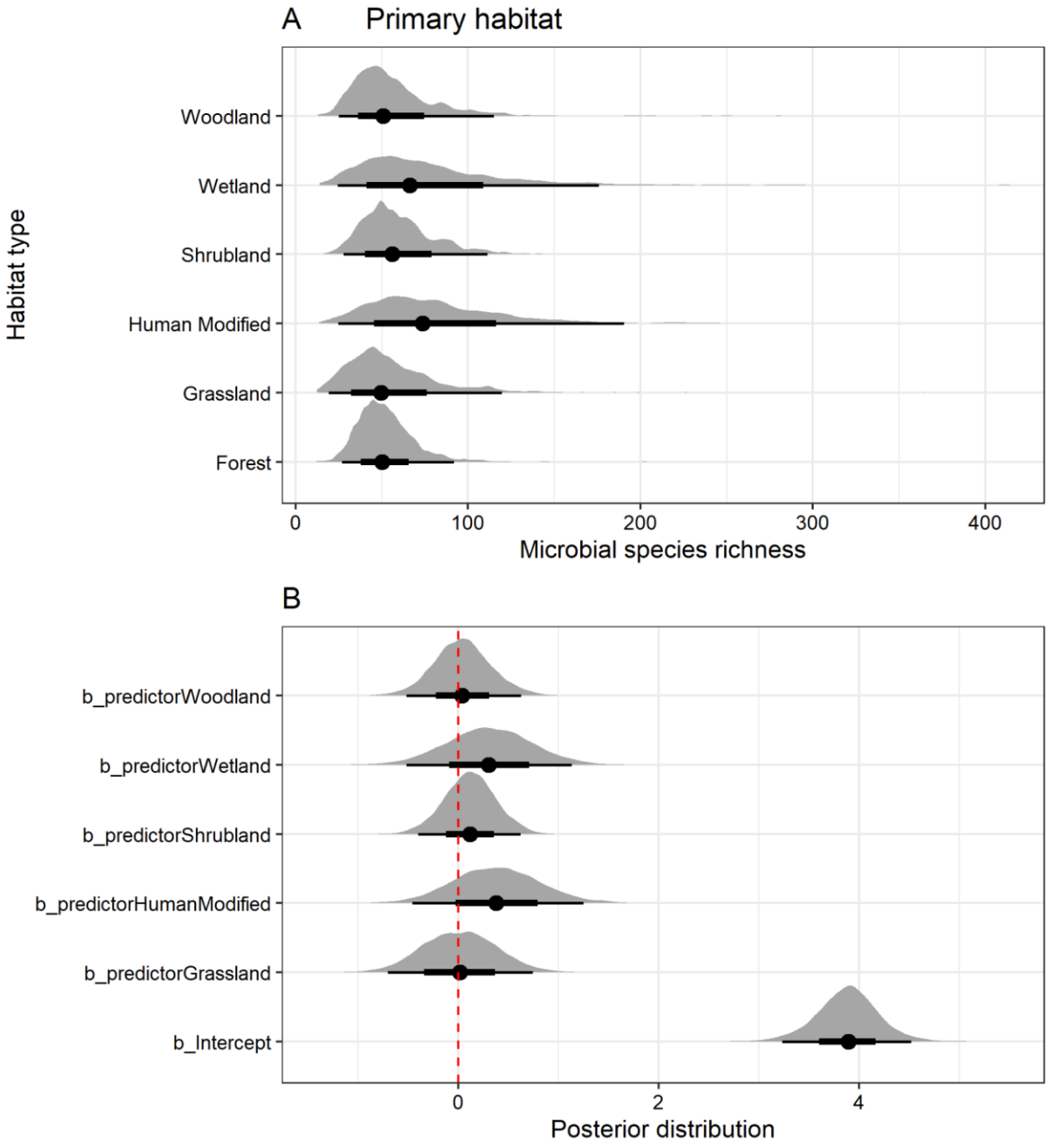

Figure S9. The results of our single regression model where alpha species richness was the response variable and body mass was the predictor variable, log10-transformed to meet modelling assumptions. Panel A represents predicted draws from our bayesian model where the red line is the average relationship and each individual line represents a separate posterior prediction, and the points represent the raw data. Panel B shows the posterior distribution of the relationship for body mass.

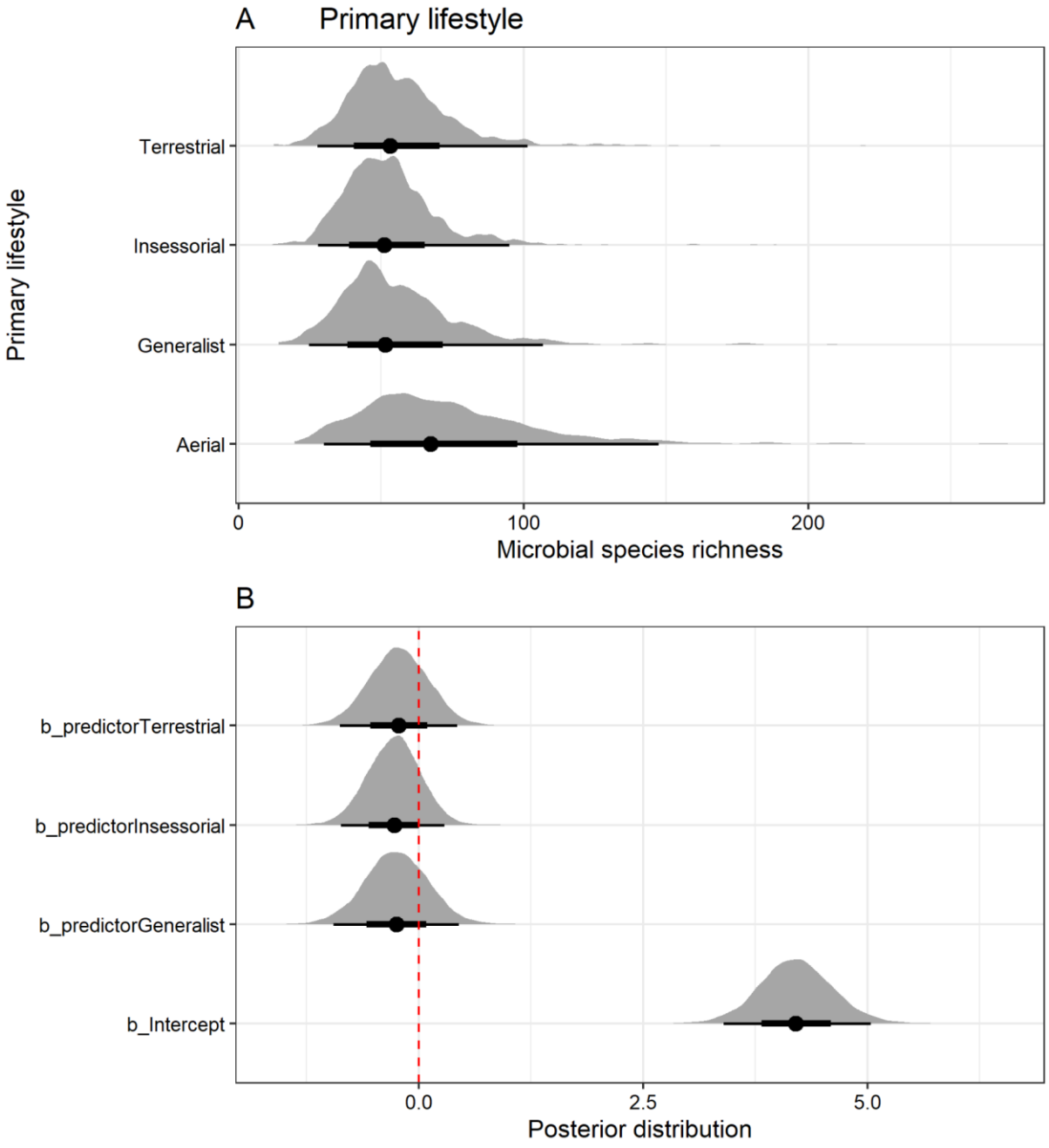

Figure S10. The results of our single regression model where alpha species richness was the response variable and body mass was the predictor variable, log10-transformed to meet modelling assumptions. Panel A represents predicted draws from our bayesian model where the red line is the average relationship and each individual line represents a separate posterior prediction, and the points represent the raw data. Panel B shows the posterior distribution of the relationship for body mass.

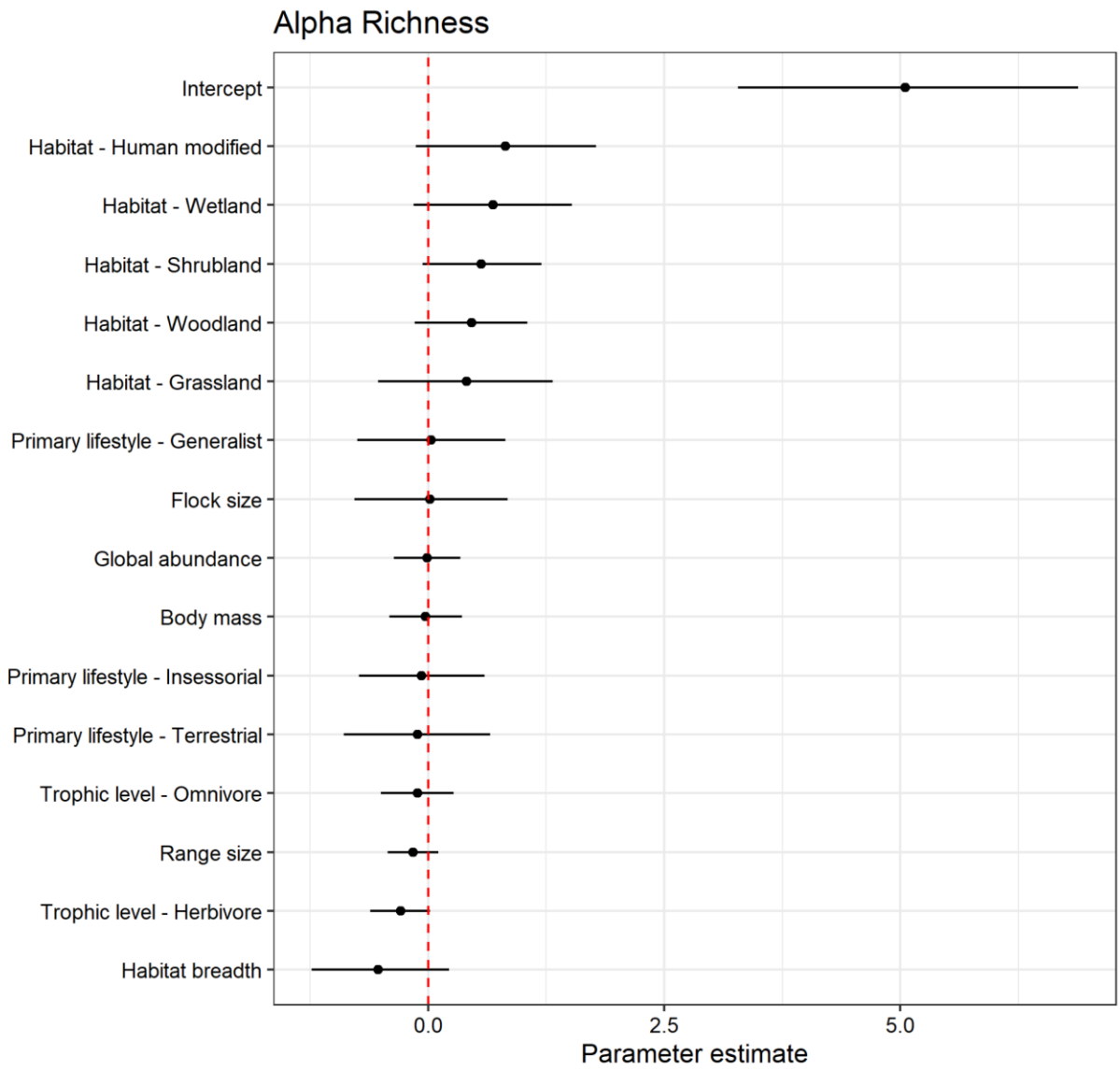

Figure S11. The results of a multiple regression model where the response variable was alpha species richness and all predictor variables were included simultaneously. The black points represent the parameter estimate and the lines represent the 95% credible interval.

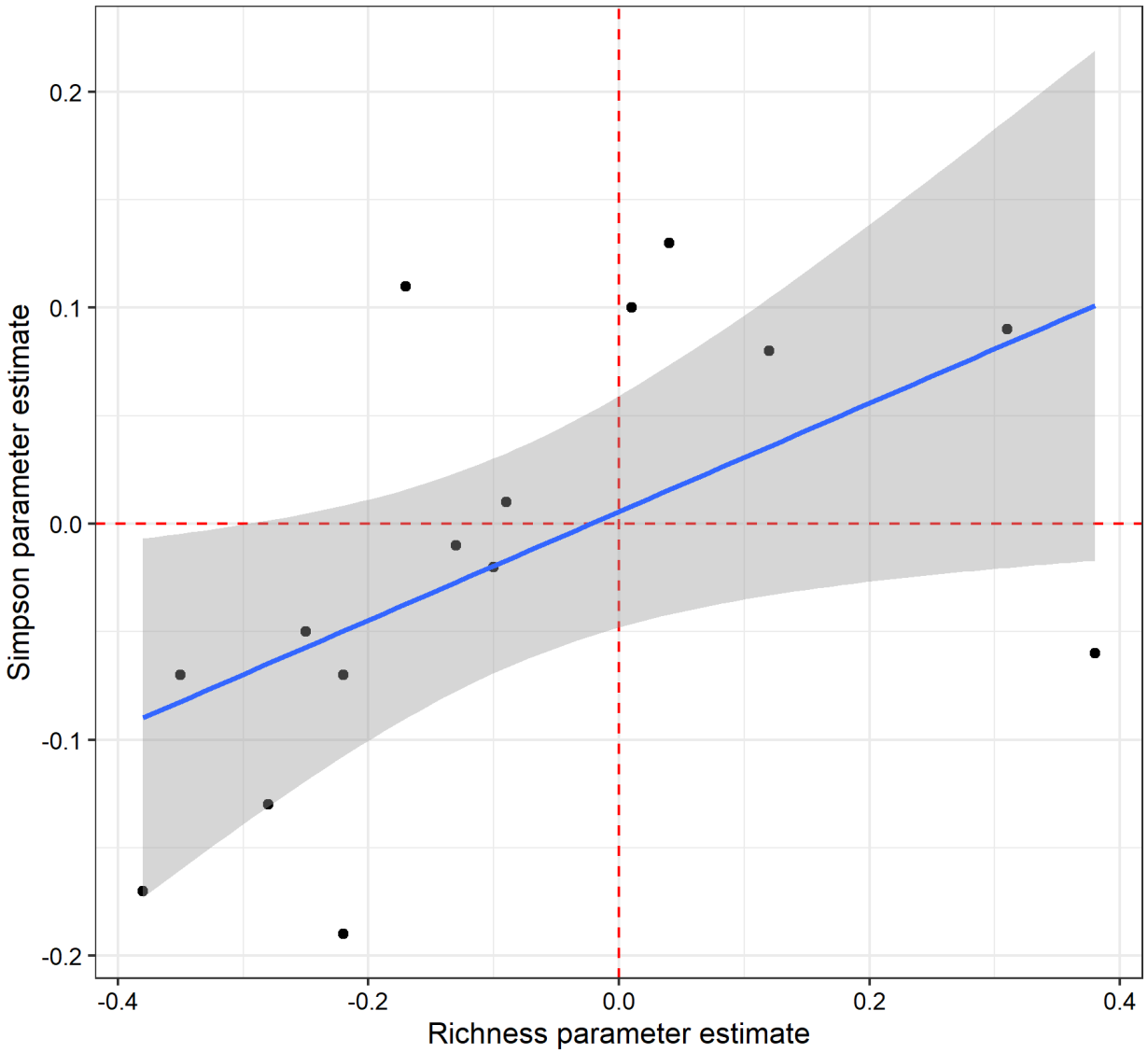

Figure S12. There was an overall strong agreement between the results of analyses where species richness was used as the response variable (x-axis) and inverse Simpson was used as the response variable (y-axis). As a result, we focused on presenting the results of species richness in the main text.

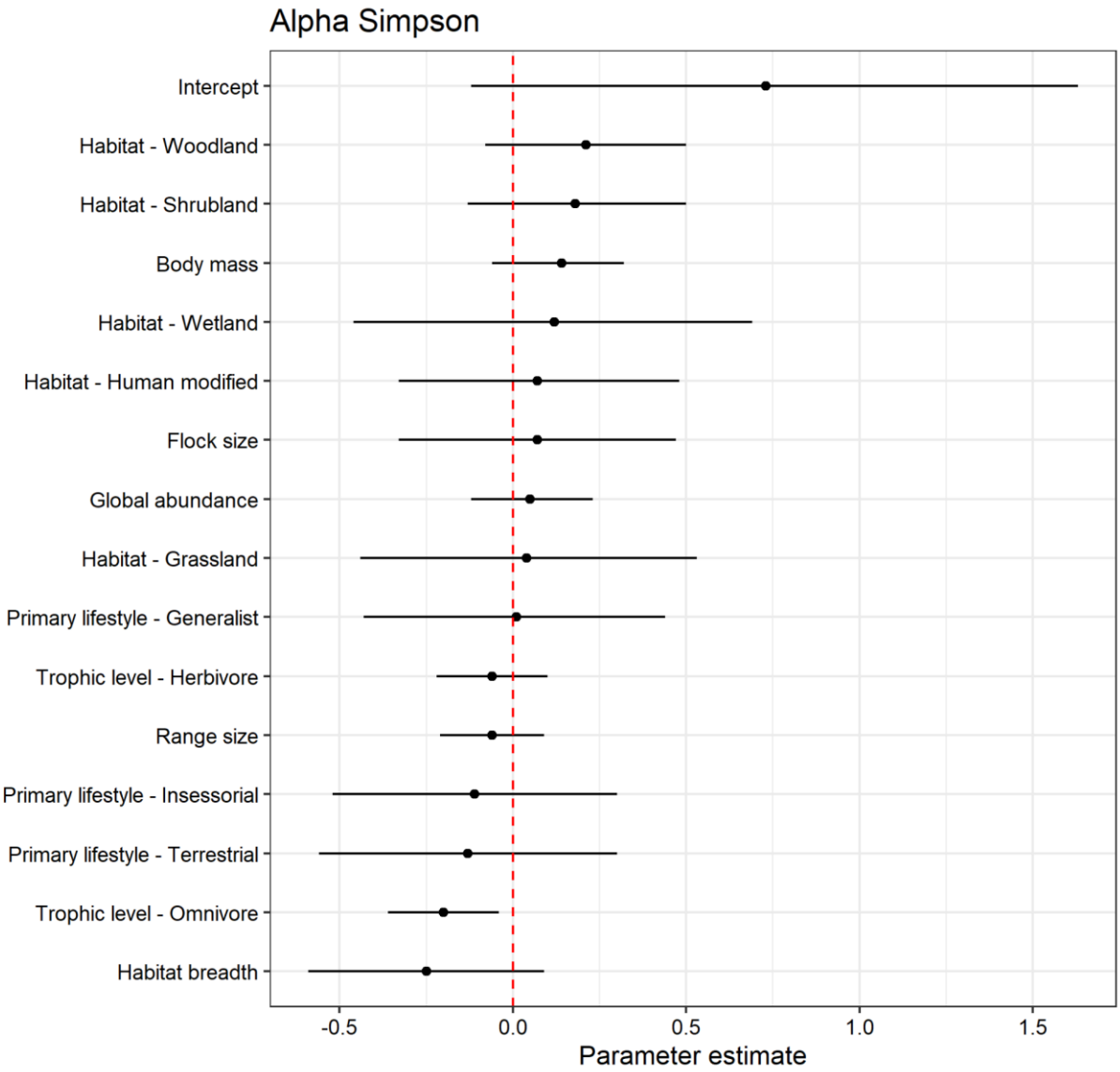

Figure S13. The results of a multiple regression model where the response variable was alpha Simpson diversity and all predictor variables were included simultaneously. The black points represent the parameter estimate and the lines represent the 95% credible interval.

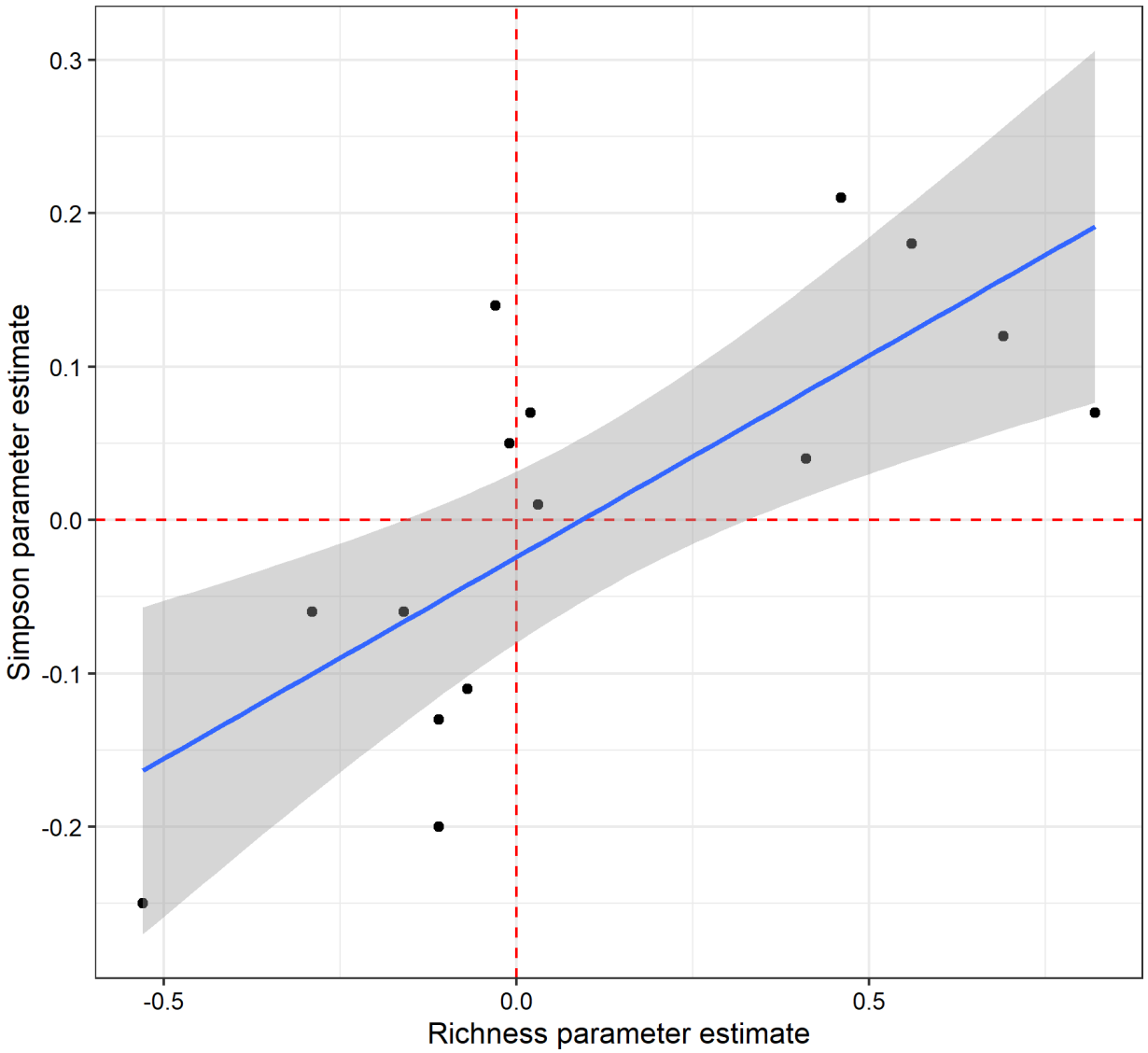

Figure S14. As with single regression results (Figure S12), there was an overall strong agreement between the results of multiple regression analyses where species richness was used as the response variable (x-axis) and inverse Simpson was used as the response variable (y-axis). As a result, we focused on presenting the results of species richness in the main text.

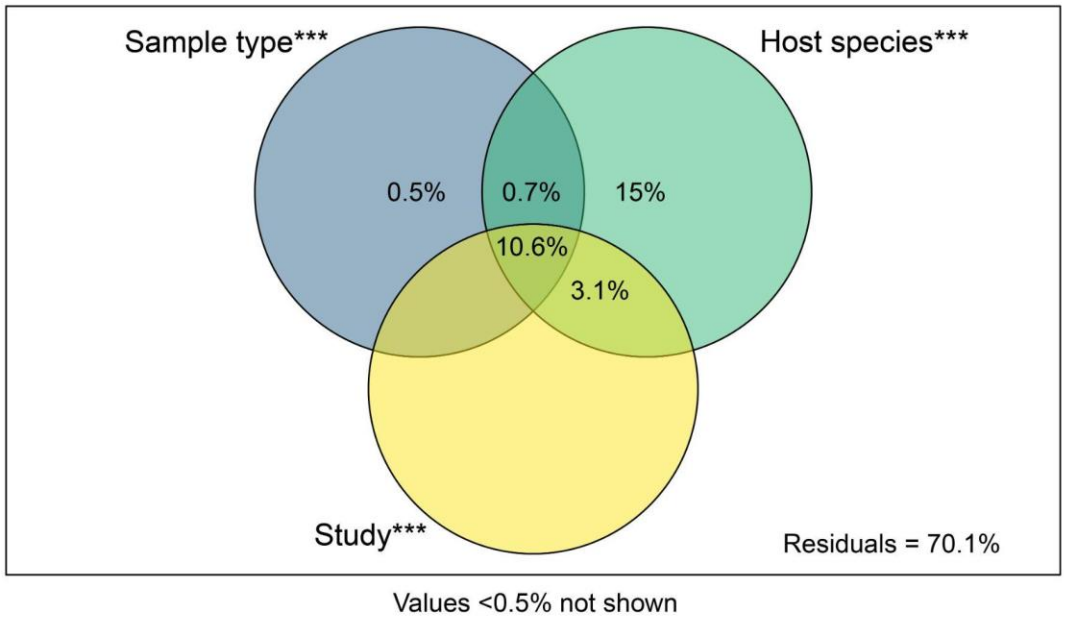

Figure S15. Variance partitioning analysis of Bray-Curtis dissimilarities focusing on the influence of technical factors on the microbiomes studied. *** p-values for permutation-based tests for the individual factors is < 0.001.