

RESEARCH ARTICLE

Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History

Andrew W. Brooks^{1,2‡}, Kevin D. Kohl^{1‡}, Robert M. Brucker^{1,3‡}, Edward J. van Opstal¹, Seth R. Bordenstein^{1,4*}

1 Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, United States of America, **2** Vanderbilt Genetics Institute, Vanderbilt University, Nashville, Tennessee, United States of America, **3** The Rowland Institute at Harvard, Harvard University, Cambridge, Massachusetts, United States of America, **4** Department of Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, Tennessee, United States of America

‡ These authors share first authorship on this work.

* s.bordenstein@vanderbilt.edu



CrossMark
click for updates

 OPEN ACCESS

Citation: Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR (2016) Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. *PLoS Biol* 14(11): e2000225. doi:10.1371/journal.pbio.2000225

Academic Editor: David Relman, Stanford University School of Medicine, United States of America

Received: June 7, 2016

Accepted: October 20, 2016

Published: November 18, 2016

Copyright: © 2016 Brooks et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All sequencing and mapping files are available from the Dryad database repository: doi:10.5061/dryad.n3v49. A GitHub repository contains custom analysis scripts and all of the necessary data for figure reconstruction (including BIOM Tables and Mapping files) for each clade are also publicly available: <https://github.com/awbrooks19/phylosymbiosis>.

Funding: Rowland Institute at Harvard University Junior Fellowship to RMB. National Science

Abstract

Phylosymbiosis was recently proposed to describe the eco-evolutionary pattern, whereby the ecological relatedness of host-associated microbial communities parallels the phylogeny of related host species. Here, we test the prevalence of phylosymbiosis and its functional significance under highly controlled conditions by characterizing the microbiota of 24 animal species from four different groups (*Peromyscus* deer mice, *Drosophila* flies, mosquitoes, and *Nasonia* wasps), and we reevaluate the phylosymbiotic relationships of seven species of wild hominids. We demonstrate three key findings. First, intraspecific microbiota variation is consistently less than interspecific microbiota variation, and microbiota-based models predict host species origin with high accuracy across the dataset. Interestingly, the age of host clade divergence positively associates with the degree of microbial community distinguishability between species within the host clades, spanning recent host speciation events (~1 million y ago) to more distantly related host genera (~108 million y ago). Second, topological congruence analyses of each group's complete phylogeny and microbiota dendrogram reveal significant degrees of phylosymbiosis, irrespective of host clade age or taxonomy. Third, consistent with selection on host-microbiota interactions driving phylosymbiosis, there are survival and performance reductions when interspecific microbiota transplants are conducted between closely related and divergent host species pairs. Overall, these findings indicate that the composition and functional effects of an animal's microbial community can be closely allied with host evolution, even across wide-ranging timescales and diverse animal systems reared under controlled conditions.

Author Summary

Studies on the assembly and function of host-microbiota symbioses are inherently complicated by the diverse effects of diet, age, sex, host genetics, and endosymbionts. Central to

Foundation—Division of Integrative Organismal Systems https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503623 (grant number 1456778). Received by SRB. National Science Foundation—Division of Biological Infrastructure <http://www.nsf.gov/div/index.jsp?div=DBI> (grant number 1400456). Received by KDK. National Institute of Health—Predoctoral Training Grant <https://researchtraining.nih.gov> (grant number 5T32GM080178). Received by AWB. National Science Foundation—Division of Environmental Biology <http://www.nsf.gov/div/index.jsp?div=DEB> (grant number 1046149). Received by SRB. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: ANOVA, analysis of variance; AVPR1A, arginine vasopressin receptor 1A; COI, cytochrome oxidase; EPE, expected predicted error; HSD, honest significant difference; OTU, operational taxonomic unit; PCoA, principal coordinates analysis; RFC, random forest classifier.

unraveling one effect from the other is an experimental framework that reduces confounders. Using common rearing conditions across four animal groups (deer mice, flies, mosquitoes, and wasps) that span recent host speciation events to more distantly related host genera, this study tests whether microbial community assembly is generally random with respect to host relatedness or "phylosymbiotic," in which the phylogeny of the host group is congruent with ecological relationships of their microbial communities. Across all four animal groups and one external dataset of great apes, we apply several statistics for analyzing congruencies and demonstrate phylosymbiosis to varying degrees in each group. Moreover, consistent with selection on host–microbiota interactions driving phylosymbiosis, transplanting interspecific microbial communities in mice significantly decreased their ability to digest food. Similarly, wasps that received transplants of microbial communities from different wasp species had lower survival than those given their own microbiota. Overall, this experimental and statistical framework shows how microbial community assembly and functionality across related species can be linked to animal evolution, health, and survival.

Introduction

A large body of literature has documented genetic and environmental influences on the composition of host-associated microbial communities [1–10]. Although environmental factors are considered to play a much larger role than host genetics and evolutionary history [11], host influences and their functional consequences are poorly elucidated and thus require systematic study across host–microbiota systems. Several outstanding questions remain regarding the nature of host effects on microbiota assembly. Are host–microbiota associations stochastically assembled, or might there be deterministic assembly mechanisms that predict these associations? How rapidly do microbiota differences form between closely related host species, and are interspecific microbiota differences prone to decay over evolutionary time? Can host-driven assembly of the microbiota be isolated from confounding variables such as diet, age, sex, and endosymbionts? If there are microbiota differences between species, are they functional in an evolutionarily informed manner, such that mismatches between host and interspecific microbiota lead to reductions in fitness or performance, particularly when interspecific microbiota transplants are conducted between older host species pairs?

If host-associated microbial communities assemble stochastically through environmental acquisition with no host-specific influence, then microbiota compositions across related host species will not differ from expectations based on random community assemblies and dispersal limitations. Therefore, in a common environment, microbiota will form independent of host species (Fig 1A), and any interspecific differences in microbiota composition would be arbitrary. In contrast, if hosts influence a sufficient amount of the composition of the microbiota, then under controlled rearing conditions, intraspecific microbial communities will structure more similarly to each other than to interspecific microbial communities (Fig 1B). Similarly, if microbial communities are randomly established or are not distinguishable with regard to host evolutionary relationships, then dendrograms illustrating beta diversity distance relationships between microbial communities will not parallel the phylogeny of the host species (Fig 1C). However, if microbial communities are distinguishable, then hosts with greater genetic divergence may exhibit more distinguishable microbiota. In this case, there will be congruence between the host phylogeny and microbiota dendrogram (Fig 1D). As this outcome is not likely due to coevolution, cospeciation, or cladogenesis of the entire microbial community from a last common ancestor, "phylosymbiosis" was proposed as a new term that

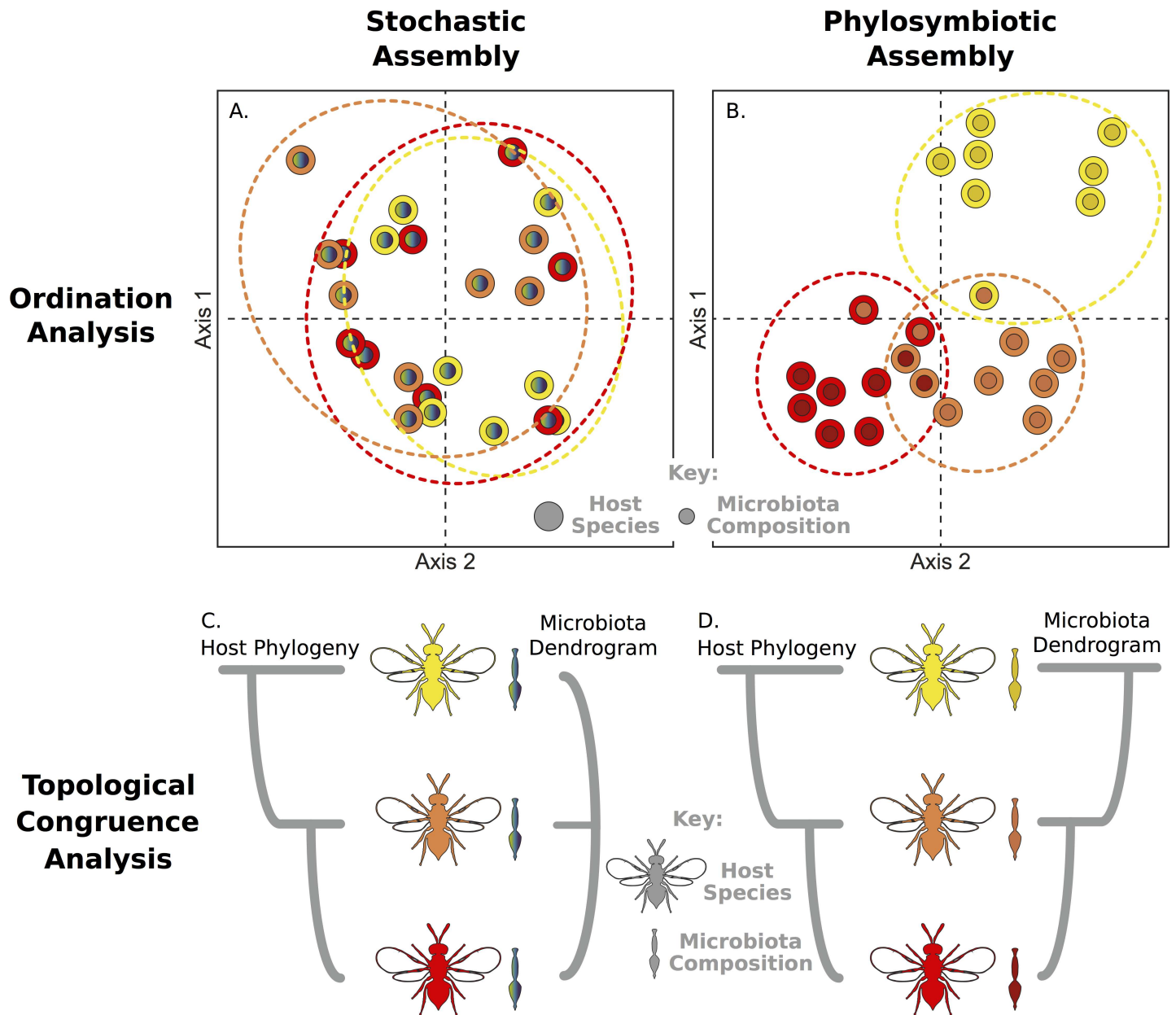


Fig 1. Analyses and predictions that can distinguish stochastic host-microbiota assembly from phylosymbiosis under controlled conditions. Two-dimensional ordination plots depict hypothetical microbiota similarity under (A) stochastic versus (B) phylosymbiotic models. Dashed lines represent host-specific clustering. Topological congruence analyses between host phylogeny (evolutionary relatedness) and microbial community dendrogram (ecological relatedness) depict the pattern expected for (C) stochastic versus (D) phylosymbiotic host-microbiota assembly.

doi:10.1371/journal.pbio.2000225.g001

does not necessarily presume that members of the microbial community are constant, stable, or vertically transmitted from generation to generation [1,12]. Rather, phylosymbiosis refers to an eco-evolutionary pattern in which evolutionary changes in the host associate with ecological changes in the microbiota.

Phylosymbiosis leads to the explicit prediction that as host nuclear genetic differences increase over time, the differences in host-associated microbial communities will also increase. Indeed, phylosymbiosis has been observed in natural populations of sponges [13], ants [10], bats [14], and apes [15,16]. However, other studies on termites [17], flies [18–20], birds [21],

and mice [22] have not observed strict patterns of phylosymbiosis or host-specific microbial signatures. In natural population studies, determining the forces driving phylosymbiosis is equivocal, as both environmental and host effects can covary and contribute to microbiota assembly. Importantly, major effects of the environment, age, or sex may overwhelm the ability to detect phylosymbiosis. Indeed, diet is a stronger determinant of whole microbial community structure than genotype in lab-bred mice [23]. Additionally, conjecture about the formation of host-specific communities should be resolved in a wider context, especially their functional significance, as microbiotas may be inconsequential to host biology or uniquely situated for certain host genotypes and fitness. Thus, the prevalence and functional significance of phylosymbiosis is uncertain and requires reductionist approaches to discriminate among the frequently confounded variables of host, environment, development, sex, and even endosymbiont status.

Here, we quantify phylosymbiosis under laboratory conditions to control for environmental and host rearing variation. Prior investigations of phylosymbiosis have not typically controlled for these confounding variables, with the exception of male *Nasonia* wasps [1,2] and *Hydra* [5,24]. Specifically, we reared 24 species in the laboratory while controlling for sex (virgin females), age, diet, and endosymbionts, thus removing major environmental variables and isolating the contribution of host species on microbiota assembly. The experimental systems, or “host clades,” span four species of *Nasonia* parasitic jewel wasps, six species of *Drosophila* fruit flies, eight species of *Anopheles*, *Aedes*, and *Culex* mosquitoes, and six species of *Peromyscus* deer mice. An externally derived dataset with seven members of the hominid lineage [16] provides another mammalian and multigenus clade for reference and facilitates examination of natural populations in which phylosymbiosis was previously documented. Together, the five host clades include 31 distinct taxa and span a range of estimated divergence times from 0.2–108 million y. Last, we test the hypothesis that phylosymbiosis represents a functional association through a series of microbial transplants with autochthonous (intraspecific) and allochthonous (interspecific) microbiota in *Nasonia* and *Peromyscus*. We expect that an experimentally mediated disruption of phylosymbiosis will have functional costs that may lower host fitness or performance in an evolutionarily informed manner. Our findings demonstrate that a consistent set of controlled experimental and bioinformatic approaches in comparative microbiota studies can isolate host-driven phylosymbiosis.

Results

Host Clade Differentiates Microbial Communities

Phylosymbiosis predicts that host clades will harbor distinguishable microbial communities (e.g., jewel wasps versus fruit flies versus deer mice, etc.) and that more closely related host clades will exhibit more similar microbial communities (e.g., insects versus mammals). Indeed, at a broad scale, we found that host clades harbored relatively distinct microbial communities (Fig 2A, ANOSIM, $R = 0.961$, $p < 1e-6$). Furthermore, there was significant microbiota differentiation between the mammalian and invertebrate host clades in the principle coordinates analysis (PCoA) (Fig 2A, ANOSIM, $R = 0.905$, $p < 1e-6$). The PCoA shows insect groups separating along two dimensions of a plane, with the mammals distinguished orthogonally from that plane in a third dimension, suggesting that variance in insect microbial communities is fundamentally different than that in mammals. As is well established, the gut communities of mammals were dominated by the bacterial classes Clostridia (Firmicutes) (Fig 2B, hominid 42%, *Peromyscus* 37%) and Bacteroidia (Bacteroidetes) (Fig 2B, hominid 15%, *Peromyscus* 37%), while the insect clades were dominated by Proteobacteria (Fig 2B, *Drosophila* 78%, mosquito 69%, *Nasonia* 77%). This same bacterial divide is also seen in the network analysis, with significant clustering of the insect microbial communities around Proteobacteria, and the

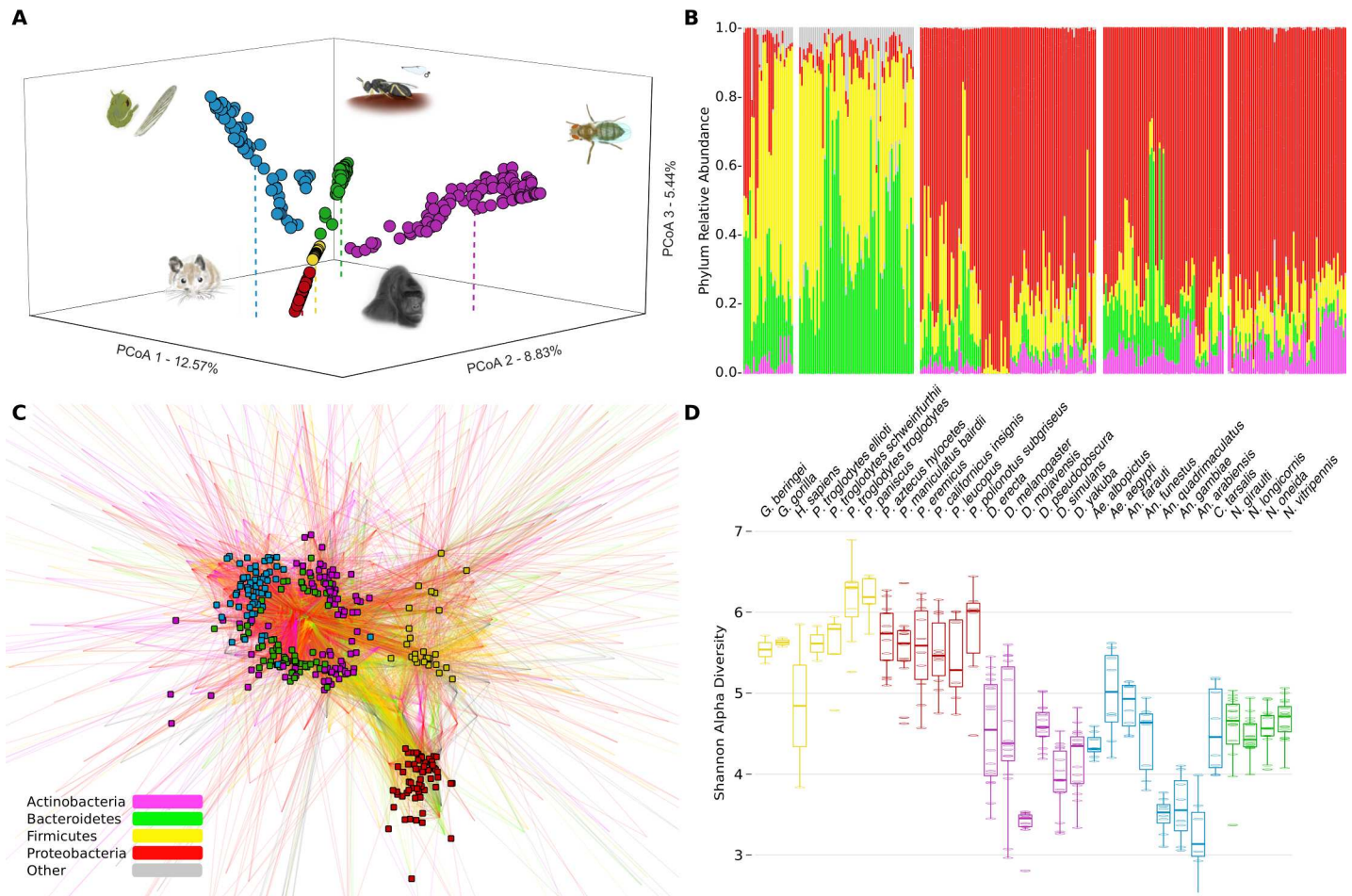


Fig 2. Meta-analysis of microbiota variation across five host clades. (A) PCoA analysis of Bray–Curtis ecological similarity in three dimensions based on 99% operational taxonomic unit (OTU) cutoff, with colors depicting clade of origin. (B) Phylum level relative abundance for all samples, with a key provided in C. (C) Network analysis in which small squares depict samples, with their color indicating clade of origin. Lines connect genus-level OTUs to samples and are weighted by occurrence and colored by OTU phylum. (D) Shannon alpha diversity for each host species. Small ellipses depict individual samples, and dark lines indicate the species' median diversity. The lower and upper end of each box represent the 25th and 75th quartiles, respectively. Whiskers denote the 1.5 interquartile range. Data available at [26] in folders (A) Fig_2A, (B) Fig_2B, (C) Fig_2C, (D) Fig_2D.

doi:10.1371/journal.pbio.2000225.g002

mammal microbial communities around subsets of shared and unique Firmicutes and Bacteroidetes (G-test, $p < 1e-6$, Fig 2C). Microbial diversity as measured by the Shannon index [25] was approximately 35% higher in mammalian hosts compared to insects, indicating more diverse symbiont communities among the mammalian clades (Fig 2D; Nested analysis of variance [ANOVA]: phylum effect [mammals versus insects]: $F_{1,302} = 419.82$, $p < 0.001$; clade effect nested within phylum: $F_{3,298} = 18.46$, $p < 0.001$; species effect nested within clade and phylum: $F_{26,272} = 7.94$, $p < 0.001$).

We implemented a random forest classifier (RFC) supervised learning algorithm to quantify the degree to which individual microbial communities can be classified into their respective host clade. RFC models show a strong ability to classify microbial communities to their correct host clades based on OTUs (98.5% classification accuracy) (S1 Table). Additionally, models distinguish mammals and insect samples with high accuracy (95.9% classification accuracy) (S1 Table). Cross-validation prevents overfitting by ensuring that classification accuracy is assessed using only samples excluded from model training. We also used RFC models

to identify the most distinguishing bacterial taxonomic level for both interclade distinction and the divide between mammals and insects. Genera provided the strongest ability to predict host clade (99.0% classification accuracy) (S1 Table); however, the major groups of insects and mammals were better distinguished by family-level community classification (98.3% classification accuracy) (S1 Table). Taken together, these results illustrate that evolutionary relationships of the host clades broadly covary with differences in microbial communities. While differentiation of the five clades could in part be attributable to varied experimental conditions for each animal group (since they were reared separately), clustering of the vertebrate microbial communities from the insect microbial communities is independent of rearing conditions and suggests a host-assisted structuring of microbial communities.

Intraspecific Microbial Communities Are Distinguishable within Host Clades

Phylosymbiosis predicts that an individual's microbial community will exhibit higher similarity to communities of the same host species than to those from different host species. The degree of similarity can be variable but should correlate with genetic relatedness of the host species. Pairwise comparisons of beta diversity distances between all individuals within each host clade reveal that the average distance between microbial communities within a species is always less than between species (S1 Fig). Summarized beta diversity also reveal lower intraspecific versus interspecific distances, with significant differences observed for all clades (Fig 3A, Each dataset: Mann–Whitney U, $p < 1e-6$).

We next evaluated intraspecific microbiota clustering through Bray–Curtis beta diversity interrelationships with PCoA and statistically assessed the strength of interspecific microbiota distinguishability with ANOSIM (Fig 3B). Visualization of the first three principle components revealed that individual samples clustered around their respective species' centroid position. In all host clades, each host species harbored significantly distinguishable microbial communities (Fig 3B, ANOSIM $p < 0.001$ for all host clades). Notably, the ANOSIM R-values of interspecific microbiota distinguishability within a host clade positively correlated with the maximal age of divergence of the species in the host clades (Fig 3C, Regression Analysis Log Transformed Clade Age, $R^2 = 0.92$, $p = 0.006$; Untransformed Clade Age, $R^2 = 0.70$, $p = 0.048$). Thus, host clades with higher total divergence times between species had stronger degrees of microbiota distinguishability, while less diverged host clades exhibited less microbiota distinguishability. For example, with an estimated host divergence time of 108 million y [27], mosquitoes showed the greatest distinguishability of their microbiota. Conversely, in *Nasonia* jewel wasps, which only diverged between 200,000 and 1 million y ago [28], the relative strength of clustering was less distinct but still statistically significant. The three intermediate aged clades showed corresponding intermediate levels of clustering: *Drosophila* had an estimated divergence time of 62.9 million y [29], hominids diverged 9 million y ago [30], and *Peromyscus* diverged 11.7 million y ago [31]. Therefore, the phylosymbiotic prediction that host species will exhibit significant degrees of specific microbiota assembly was supported in these observations, even under highly controlled conditions in the laboratory models. Microbiota specificity was maintained among very closely related and very divergent species, and a connection was observed between the magnitude of host genetic divergence and microbiota similarity.

Supervised Classification: Microbiota Composition Predicts Host Species

As microbiota clustering was supported within species across all five animal clades, it should be possible to model the strength of how well communities of bacteria predict their host

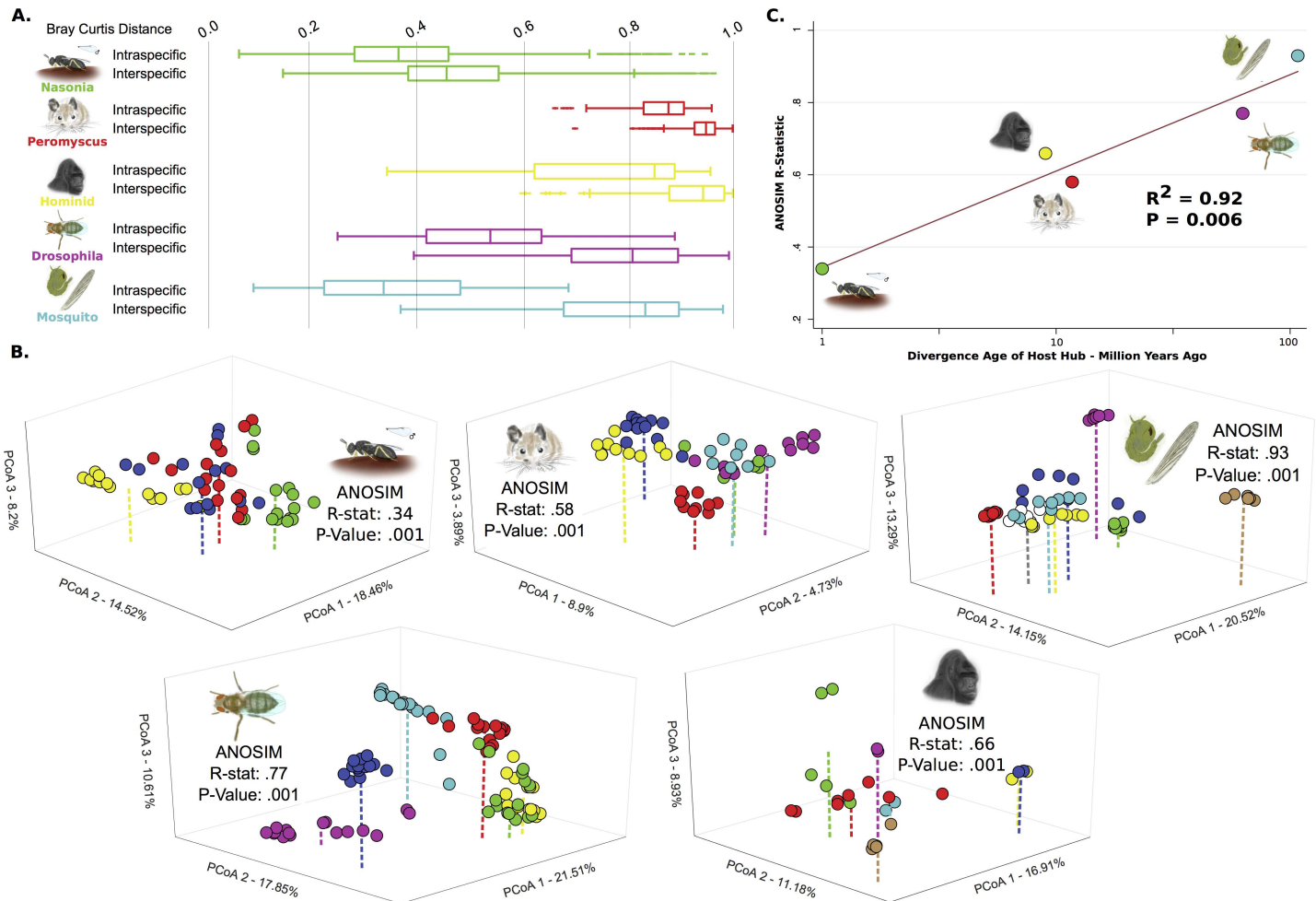


Fig 3. Intraspecific versus interspecific microbial community variation within and between host clades. (A) Box-and-whisker plot of intraspecific and interspecific Bray–Curtis distances between samples for each clade. Boxes represent the 25th to 75th quartiles, with the central line depicting the group median and whiskers showing the 1.5 interquartile extent. (B) PCoA of Bray–Curtis distances with first three most distinguishing dimensions shown. Colors represent different species and correspond to the colors in Fig 4. (C) Regression analysis measuring the correlation between the evolutionary age of host clade divergence on a log scale and the ANOSIM R-values of intraspecific microbiota distinguishability from part B for each host clade. Data available at [26] in folders (A) Fig_3A_&_S1, (B) Fig_3B, (C) Fig_3C.

doi:10.1371/journal.pbio.2000225.g003

species and how specific members of the microbiota affect these predictions. We therefore used RFC models trained on the microbiota of each host clade to evaluate classification accuracy (i.e., the percentage of assigning microbiota to their correct host species) and the expected predicted error (EPE, i.e., the ratio of model accuracy relative to random classification). RFC results indicated that the operational taxonomic units (OTUs) for *Drosophila* and *Peromyscus* and genus taxonomic levels for hominid, mosquito and *Nasonia* have the highest classification accuracies, with significant EPE observed for all clades (EPE > 2, S1 Table). At the genus level, the mosquito and *Drosophila* host clades exhibited the strongest results (mosquito, classification accuracy = 99.8%, EPE = 558.9; *Drosophila*, classification accuracy = 97.2%, EPE = 31.7). Other host clades demonstrated significant but comparatively lower strength models. The reduced predictive power of these models may be due to a number of factors, such as a lower number of host species (*Nasonia*, classification accuracy = 88.7%, EPE = 13.4), uneven sample representation from each species (hominid, classification accuracy = 53.4%, EPE = 2.1), and lower sequencing coverage (*Peromyscus*, classification accuracy = 61.4%, EPE = 2.5).

To determine the most distinguishing genera of the bacterial community, we examined the resulting loss of model classification accuracy when each genus was excluded from RFCs (S2 Table). Distinguishability within the *Drosophila*, *Nasonia*, and mosquito clades was driven primarily by genera in Proteobacteria, which represent five (14.0% model accuracy), seven (11.3% model accuracy), and eight (18.2% model accuracy) of the top ten genera, respectively. Three of the ten most distinguishing genera in *Drosophila* females are from the Acetobacteraceae family (9.5% model accuracy), previously recognized to be “core” microbiota members [19,32]. Three of the twenty most distinguishing genera in *Nasonia* females were closely related symbionts from the Enterobacteriaceae family (genera: *Proteus*, *Providencia*, *Morganella*; 3.1% model accuracy), consistently found in our previous studies of *Nasonia* males [1,2]. Eight genera from the phylum Proteobacteria dominate mosquito female distinguishability, primarily three Gammaproteobacteria of the order Pseudomonadales (8.2% model accuracy), and three Betaproteobacteria of the family Comamonadaceae (5.9% model accuracy). Hominid interspecific distinguishability was driven by the phylum Firmicutes, particularly of the order Clostridiales that contains three of the most distinguishing genera (1.5% model accuracy). The genus *Allobaculum* conferred nearly double the distinguishing power of any other bacteria in *Peromyscus* (3.8% model accuracy), and it is associated with low-fat diet, obesity, and insulin resistance in mice [33]. As may be expected, genera of the abundant phyla Firmicutes and Bacteroidetes dominated the majority of distinguishability in *Peromyscus* (10.6% model accuracy), but genera from Proteobacteria in the family Helicobacteraceae comprised four of the top eleven genera (4.4% model accuracy). Overall, microbiota composition can be used to predict host species with high accuracy, and genera commonly observed in other studies of these host clades underlie interspecific distinguishability.

Phylosymbiosis Is Common within Host Clades

The major prediction of phylosymbiosis is that phylogenetic relatedness will correlate with beta diversity relationships of microbial communities among related host species. Microbiota dendrograms were constructed by collapsing individual samples to generate an aggregate microbial community for each species and then by comparing relationships of their beta diversity metrics. The matching cluster and Robinson–Foulds tree metrics were utilized to calculate host phylogenetic and microbiota dendrogram topological similarity, with normalized distances ranging from 0.0 (complete congruence) to 1.0 (complete incongruence; [34]). Matching cluster weights topological congruency of trees, similar to the widely used Robinson–Foulds metric [34,35]. However, matching cluster takes into account sections of subtree congruence and therefore is a more refined evaluation of small topological changes that affect incongruence. Significance of the matching cluster and Robinson–Foulds analyses was determined by the probability of randomized bifurcating dendrogram topologies yielding equivalent or more congruent phylosymbiotic patterns than the microbiota dendrogram. Additionally, using the same methodology, matching cluster and Robinson–Foulds metrics were evaluated for Bray–Curtis, unweighted UniFrac [36], and weighted UniFrac [36] beta diversity dendrograms at both 99% and 97% clustered OTUs (S2 Fig). The cytochrome oxidase I (COI) gene was used to construct the phylogeny for each host clade, which compared well to established phylogenetic or phylogenomic trees for all species included in the study (*Nasonia* [27]; *Drosophila* [28]; hominids [29]; mosquitoes [26]). *Peromyscus* was further resolved with an additional marker (arginine vasopressin receptor 1A [AVPR1A]) to reflect the latest phylogenetic estimates [37,38].

Nasonia female wasps exhibited an equivalent phylogenetic tree and microbial community dendrogram, representing exact phylosymbiosis (*Nasonia* wasps, Fig 4A). These results parallel

previous findings in *Nasonia* males [1,2]. Despite congruency, the *Nasonia* clade has limited topological complexity with only four species, therefore resulting in a relatively marginal significance. Mice also show nearly perfect congruence, with the exception of *Peromyscus eremicus* (Fig 4B). *Drosophila* fruit flies (Fig 4C) showed the lowest topological congruency but were still moderately significant. Four of the six species show correct topological relationships, while the microbial community relationships of *Drosophila pseudoobscura* and *D. erecta* are topologically swapped. These results are different from previous findings in *Drosophila* that utilized a different experimental design, set of taxa, and sequencing technology [19]. However, the evidence for phylosymbiosis is tentative in *Drosophila* as, unlike other clades, there is no significant congruence for either unweighted or weighted UniFrac metrics (S2 Fig). Previous studies detected no pattern of phylosymbiosis across *Drosophila* species [19], which could be attributed to *Drosophila*'s constant replenishment of microbes from the environment [18,20] or the dominance by the bacterial genus *Acetobacter*, which is important for proper immune and metabolic development [19]. The two additional clades, mosquitoes and hominids, showed significant phylosymbiosis (Fig 4D and 4E). Specifically, the mosquitoes showed accurate separation of *Culex* and *Aedes* genera from *Anopheles*, and the topological departures from phylosymbiosis appeared in two of the bifurcations between closely related species. The hominid microbial community dendrogram reflects the correct branching of *Gorilla* from *Homo sapiens*, followed by bonobos and chimpanzees, with the exception that one of the chimpanzee subspecies grouped more closely with the bonobo lineage. These results are similar to

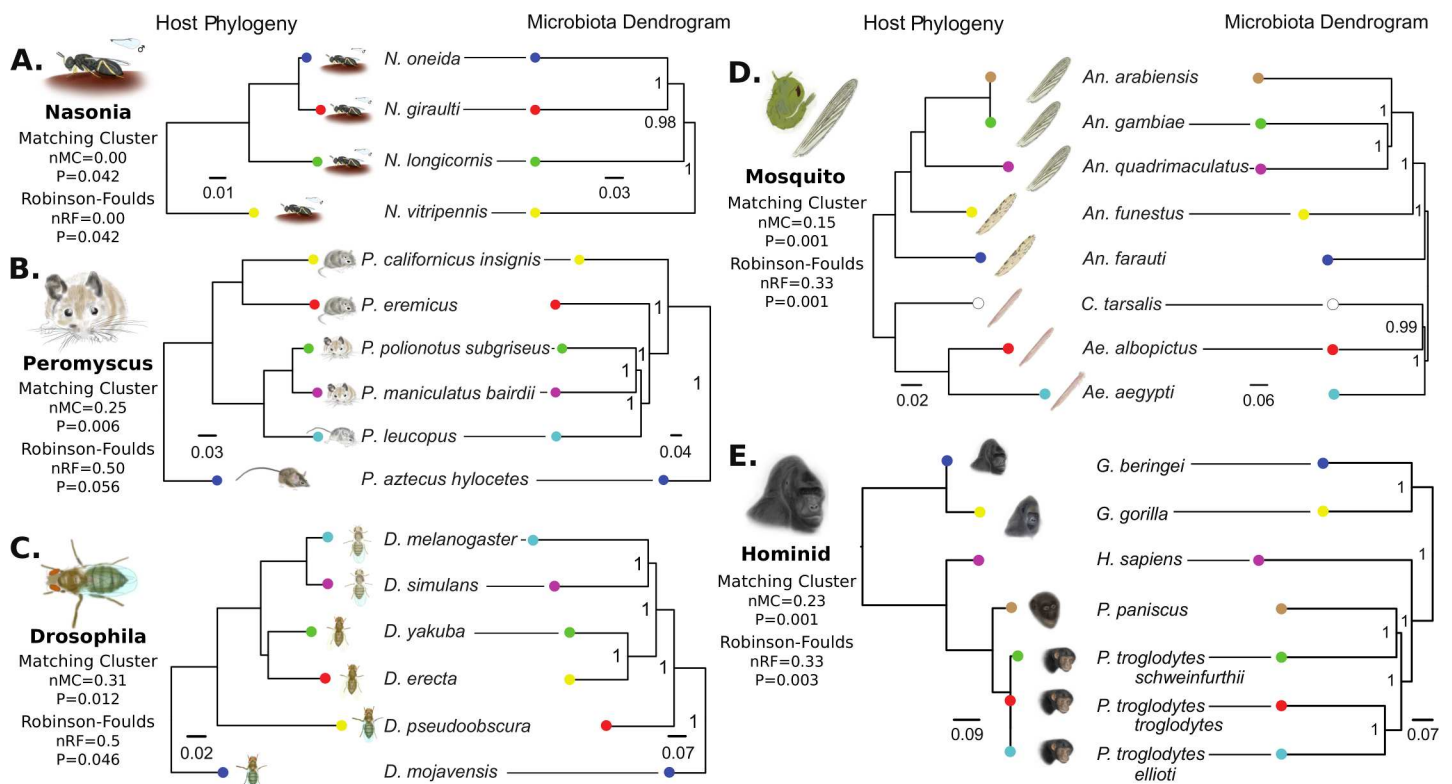


Fig 4. Phylosymbiosis between host phylogeny and microbiota dendrogram relationships. Topological congruencies are quantified by the normalized Robinson–Foulds (RF) metric, which takes into account symmetry in rooted tree shape on a scale from 0 (complete congruence) to 1 (incomplete incongruence). The normalized matching cluster (MC) metric is a refined version of the RF metric that sensitively accounts for incongruences between closely related branches. Horizontal lines connect species whose position is concordant between host phylogeny and microbiota dendrogram based on 99% OTU cutoffs, therefore requiring no topological shift to demonstrate phylosymbiosis. Data available at [26] in folder Fig_4.

doi:10.1371/journal.pbio.2000225.g004

previous observations that the relationships of the microbial communities parallel those in the host phylogeny [16]. With the exception of *Drosophila*, which yielded variable evidence for host–microbiota congruence, significant degrees of phylosymbiosis were observed across clades with varying tree similarity metrics and microbiota beta diversity analyses.

Phylosymbiosis Represents a Functional Association

Microbiota–host distinguishability and topological congruence does not strictly imply that the phylosymbiotic associations are fitness directed, though it naturally follows that a particular host species may be more ideally suited for an autochthonous versus allochthonous microbiota. We therefore performed a series of microbial transplants to test the prediction that inoculated microbiota from a different species would decrease aspects of host performance or fitness in contrast to inoculated microbiota from the same species. Moreover, if there is selection on host–microbiota interactions such that microbiotas are uniquely or better situated for resident host backgrounds, then transplanted microbiota from a divergent species could drive more pronounced reductions in host functions than transplanted microbiota from a closely related species.

In *Peromyscus*, we followed a previously established protocol [39] to transplant the microbial communities from six rodent donor species into a single recipient species, *P. polionotus*, as well as a control group in which the microbial communities from *P. polionotus* were introduced to intraspecific individuals of *P. polionotus*. Inventories of fecal microbiota from donor and recipient mice revealed that portions of the donor microbiota successfully transferred. The estimated amount of transplanted OTUs and their relative abundance ranged from 6.5%–26.2% and 11.4%–40.7%, respectively, when analyzed at the 99% OTU cutoff level. Variation in the transfer of foreign microbes was dependent on donor species and its divergence from the recipient species (S3 Fig). We then measured dry matter digestibility, or the proportion of food material that is digested by the animal. Consistent with selection on host–microbiota interactions, mice that were inoculated with microbial communities from more distantly related hosts exhibited decreased dry matter digestibility (Fig 5). These results were only significant when the group receiving feces from *P. eremicus* donors was removed (Fig 5). Notably, the microbiota of *P. eremicus* is not congruent with our predictions of phylosymbiosis (Fig 4). Thus, only the taxa showing phylosymbiosis exhibited the functional trend with digestibility. Distantly related donor species (*Neotoma lepida* and *Mus musculus*) did not drive significance, as the correlation remained statistically significant when investigating only *Peromyscus* donors (excluding *P. eremicus*; Fig 5).

In the most extreme cases in which mice were inoculated with the microbial communities from *P. californicus* or *M. musculus*, there was approximately a 3% decrease in dry matter digestibility, which is on par with the decrease in digestibility observed as a result of helminth infections in *Peromyscus* [40]. Animals must consume more food to meet energy demands when faced with decreases in digestibility. Indeed, mice inoculated with microbial communities from *P. californicus* or *M. musculus* exhibited significantly higher food intakes than the control group (S4 Fig; Tukey's honest significant difference (HSD) test: $p = 0.001$ for *P. californicus* to *P. polionotus*; $p = 0.044$ for *M. musculus* to *P. polionotus*). The mice inoculated with the microbes from *P. eremicus* performed just as well, if not better, than the control groups in terms of dry matter digestibility (Fig 5) but still had slightly higher food intakes (S4 Fig).

In *Nasonia*, we used an *in vitro* rearing system to transplant heat-killed microbial communities from three *Nasonia* donor species into larvae of *N. vitripennis* or *N. giraulti* [41]. We then measured the survival of the recipients from first instar larva to adulthood. In both *N. vitripennis* and *N. giraulti* hosts, interspecific microbiota transplantations exhibited significant

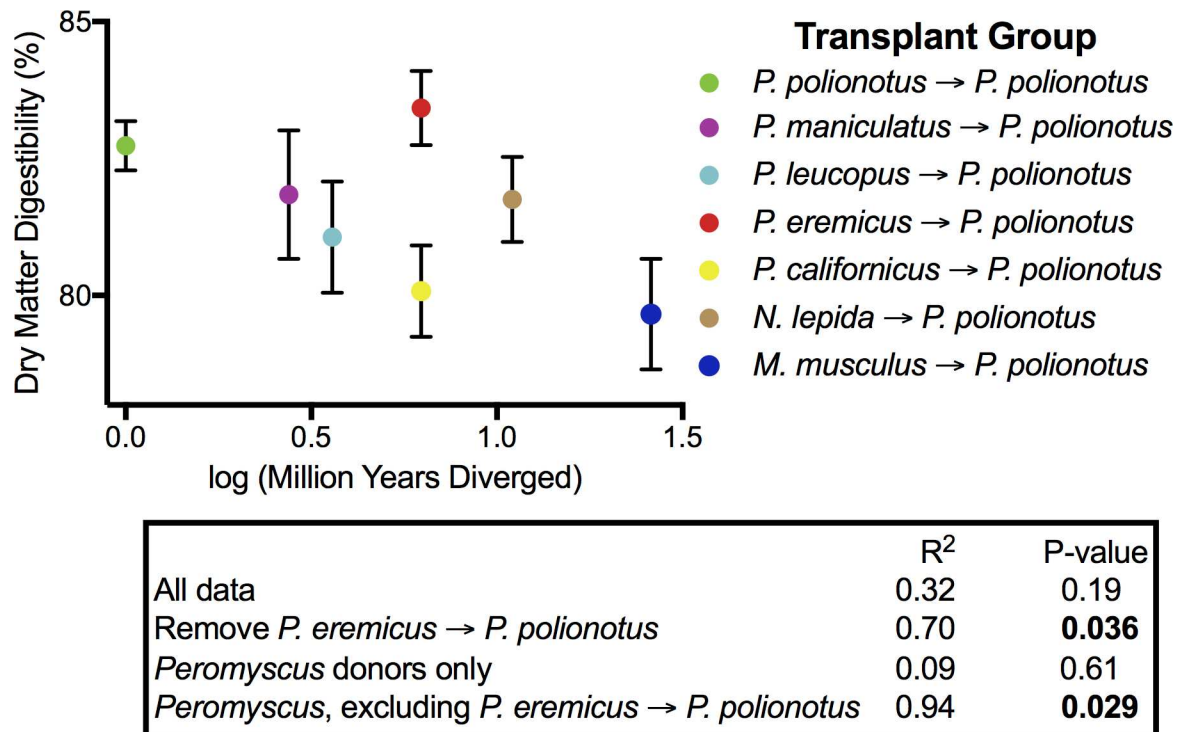


Fig 5. Effects of allochthonous and autochthonous microbial communities on the digestive performance of recipient mice. Dry matter digestibility is calculated as (g dry food ingested–g dry feces produced) / g dry food ingested. Divergence times between *P. polionotus* and donor species were determined from previously published phylogenies [37,38]. Points represent mean values ± standard error for each group ($n = 5–6$ recipients per group). Data available at [26] in folder Fig_5_&_S4.

doi:10.1371/journal.pbio.2000225.g005

decreases in survival to adulthood when compared to intraspecific microbial transplantations (Fig 6). Specifically, *N. giraulti* with a *N. vitripennis* microbiota yielded a 24.5% average survival decrease in comparison to a *N. giraulti* microbiota (Fig 6A, Mann–Whitney U, $p = 0.037$). Interestingly, *N. giraulti* with a microbiota from the more closely related *N. longicornis* exhibited a similar but nonsignificant survival reduction (23.7%, Fig 6A, Mann–Whitney U, $p = 0.086$). *N. vitripennis* with a *N. giraulti* or *N. longicornis* microbiota exhibited a 42.6% (Fig 6B, Mann–Whitney U, $p < 0.0001$) and 23.3% (Fig 6B, Mann–Whitney U, $p = 0.003$) average survival decrease in comparison to a *N. vitripennis* microbiota, respectively (Fig 6A, Mann–Whitney U, $p < 0.0001$). Comparisons were also made between noninoculated hosts and those inoculated with interspecific backgrounds (*N. giraulti* background: *N. vitripennis* inoculum $p = 0.07$, *N. longicornis* inoculum $p = 0.26$; *N. vitripennis* background: *N. giraulti* inoculum $p = 0.001$, *N. longicornis* inoculum $p = 0.15$).

Discussion

Under phylosymbiosis, host-associated microbial communities form, in part, as a result of interactions with the host rather than through purely stochastic processes associated with the environment. Specifically, we predicted that given closely related animals reared in controlled environments, the relationships of the microbiota would be congruent with the evolutionary relationships of the host species. Previous evidence for phylosymbiosis under controlled regimes existed in *Nasonia* [1,2] and *Hydra* [24], and wild populations of sponges [13], ants [10], and apes [15,16] also exhibited this pattern. Here, in a comprehensive analysis of

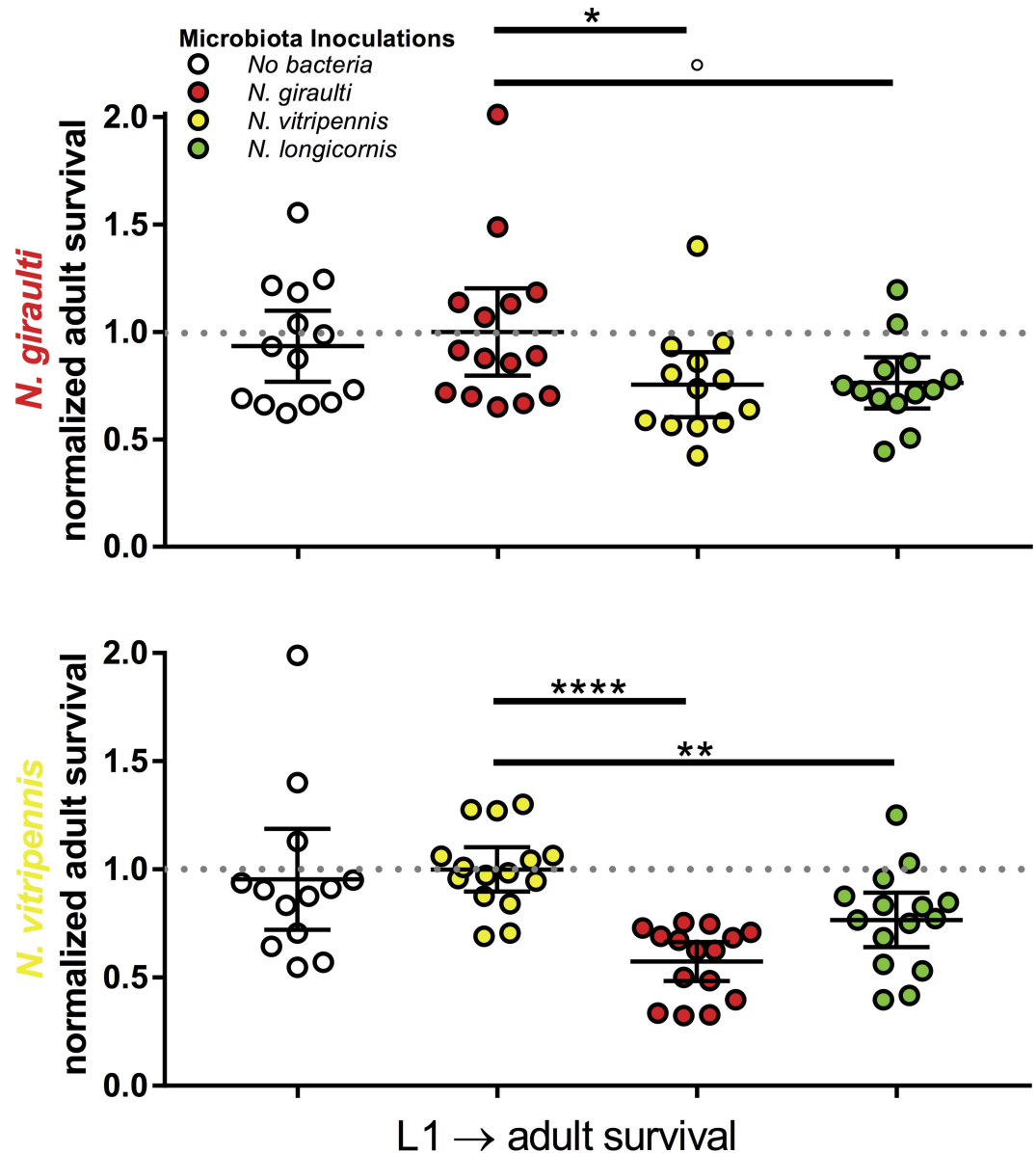


Fig 6. Effects of allochthonous and autochthonous microbial communities on the survival of *Nasonia* wasps. (A) Normalized larval-to-adult survival of *N. giraulti* wasps harboring no, self, or foreign microbiota. (B) Normalized larval-to-adult survival of *N. vitripennis* wasps harboring no, self, or foreign microbiota. Adult survival is calculated as number of adults in a transwell / number of first instar larvae in a transwell. Adult survival was normalized to the average survival of the autochthonous microbiota transplantation. Circles represent individual transwell samples, and the dashed line represents the average survival of the autochthonous microbiota transplantation normalized to 1; error bars represent 95% confidence intervals. Mann–Whitney U statistics, $^{\circ}p < 0.1$, $^*p < 0.05$, $^{**}p < 0.01$, and $^{****}p < 0.0001$. Data available at [26] in folder Fig_6.

doi:10.1371/journal.pbio.2000225.g006

phylosymbiosis in a diverse range of model systems, we report the widespread occurrence of this pattern under strictly controlled conditions as well as a functional basis in the context of host digestive performance in mice and survival in wasps. These results represent the first evidence for phylosymbiosis in *Peromyscus* deer mice, *Drosophila* flies, a variety of mosquito species spanning three genera, and *Nasonia* wasp females with the inclusion of *N. oneida*. Previous studies in *Nasonia* measured male phylosymbiosis and did not include *N. oneida* [1,2].

By rearing closely related species from the same host clade in a common environment, and by controlling age, developmental stage, endosymbiont status, and sex, the experiments rule out confounding variables that can influence microbiota relationships in comparative analyses. Eliminating these variables is important because they often substantially correlate with inter-specific differences. Thus, our findings demonstrate that a uniform experimental and bioinformatic methodology can excavate host effects on phylosymbiosis from other potentially confounding variables in comparative microbiota studies.

We observed marked differences in microbial diversity and community structure between mammalian and invertebrate host clades. Mammalian communities were more diverse and dominated by Bacteroidetes and Firmicutes, while insect-associated communities were less diverse and primarily dominated by Proteobacteria. These results are consistent with previous microbial inventories conducted in mammals and insects [6,42]. Together, these findings suggest large-scale differences in the host–microbiota interactions between mammals and insects. These differences across host phyla could be due to a variety of possibilities, including host genetics, diet, age, and rearing environment.

To remove confounding variables that structure host–microbiota assemblages and to rigorously test phylosymbiosis, we utilized an experimental design within four host clades that isolated the effects of host evolutionary relationships from other effects (i.e., diet, age, rearing environment, sex, endosymbionts). We found that host species consistently harbored distinguishable microbiota within each host clade. Additionally, we found significant degrees of congruence between the evolutionary relationships of host species and ecological similarities in their microbial communities, which is consistent with the main hypothesis of phylosymbiosis. These results importantly expand previous evidence for this eco-evolutionary pattern and demonstrate that related hosts reared under identical conditions harbor distinguishable microbial assemblages that can be likened to microbial community markers of host evolutionary relationships. It is conceivable that recently diverged species (i.e., those younger than several hundred thousand years) would have less genetic variation and fewer differences in microbiota composition. Furthermore, divergent hosts may have vast differences in physiology that overwhelm the likelihood of observing phylosymbiosis. Surprisingly, we observed phylosymbiosis to varying degrees in all host clades, and the age of clade divergence positively correlates with the level of intraspecific microbiota distinguishability. Thus, as host species diverge over time, microbial communities become more distinct [1,12], and the limits of detecting phylosymbiosis may occur at extreme scales of incipient or ancient host divergence times.

The mechanisms by which phylosymbiosis is established requires systematic investigation. Perhaps the most apparent regulator of host–microbiota interactions is the host immune system. A previous study of phylosymbiosis in *Hydra* demonstrated that antimicrobial peptides of the innate immune system are strong dictators of community composition, and expression of antimicrobial peptides are necessary for the formation of host-specific microbiota [5]. Furthermore, genome-wide association studies in humans [43], mice [8], and *Drosophila* [44] have identified a large immune effect in which host immune genes can explain variation in microbial community structure. Interestingly, host immune genes often exhibit rapid evolution and positive selection compared to genes with other functions [45,46]. While this trend is often explained by the host–pathogen arms race [45], it is also likely due to host evolutionary responses for recruiting and tending a much larger collection of nonpathogenic microbes.

Other host pathways may also underlie the observed species-specific microbiota signatures. Hosts produce glycans and mucins on the gut lining that may serve as biomolecular regulators of microbial communities [47,48]. For example, knocking out the gene for α 1–2 fucosyltransferase inhibits production of fucosylated host glycans on the gut surface and significantly alters microbial community structure [49]. Additional knockout studies have demonstrated the roles

of circadian clock genes [50], microRNAs [51], and digestive enzymes [52] in determining microbial community structure. These various physiological systems might also interact with one another and may have even evolved in tandem to regulate microbial community structure.

Alternatively, rather than hosts “controlling” their microbiota, microbes may be active in selecting which host niches to colonize. For example, hosts have been compared to ecological islands, where environmental selection of the microbiota through niche availability may occur [53]. However, given the large number of studies that demonstrate the role of microbes in improving host performance [54], we find it unlikely that hosts would assume a solely passive role in these interactions. An elegant study allowed microbial communities from various environments (soil, termite gut, human gut, mouse gut, etc.) to compete within the mouse gut [55]. This study found that a foreign community of the human gut microbiota exhibited an early competitive advantage and colonized the mouse gut first. Later, the mouse gut microbiota dominated and outcompeted the human gut microbiota [55]. Thus, community assembly is not a monolithic process of host control but likely a pluralistic combination of host control, microbial control, and microbe–microbe competition. In this context, both population genetic heritability and community heritability measurements of the microbiota will be useful in prescribing the varied genetic influences of a foundational host species on microbiota assembly [56].

The acquisition route of microbes could also influence our understanding of phylosymbiosis. If phylosymbiosis is observed when the microbiota is acquired horizontally from other hosts, the environment, or some combination of the two, then phylosymbiosis is presumably influenced by host-encoded traits such as control of or susceptibility to microbes. However, maternal transmission of microbes is argued to be a common trend in animals [57]. For example, sponges exhibit vertical transmission of a diverse set of microbes in embryos [58]. Transmission of full microbial communities is unlikely in most systems, given that the communities of developing animals tend to exhibit markedly lower diversity and distinct community structure compared to adults [1,59,60]. Thus, it is improbable that phylosymbiotic relationships are explained simply by community drift over host evolutionary divergence. There could be a subset of microbial taxa that are more likely to be transmitted from mother to offspring that in turn affect what other microbes colonize. For instance, in humans, the family *Christensenellaceae* is situated as a hub in a co-occurrence network containing several other gut microbes and has a significant population genetic heritability [61]. When *Christensenella minuta* was introduced into the guts of humanized mice, the microbial community structure was significantly altered [61]. This microbe, as well as others, can therefore be likened to a keystone taxa or “microbial hub” that can impact community structure despite low abundance [61–63]. Thus, one could hypothesize that phylosymbiotic relationships in some systems may be driven by host transmission of microbial hubs that determine whole community structure through ensuing microbe–microbe interactions. However, further work is needed to test this hypothesis.

The congruent relationships between hosts and associated microbial communities are likely maintained through their positive effects on host performance and fitness but could be neutral or harmful as well. While the importance and specificity of hosts and microbes in bipartite associations has been demonstrated on host performance [64], it is unclear whether such effects commonly occur for hosts and their complex microbial communities. If they exist, disruption of phylosymbiosis via hybridization or microbiota transplants should lead to reduced fitness or performance. For instance, hybridization experiments demonstrate negative interactions or “hybrid breakdown” between host genetics and the gut microbiota that drives intestinal pathology in house mice [65] and severe larval lethality between *N. vitripennis* and *N. giraulti* wasps [2]. Furthermore, transplant experiments show that all microbes are not equal for the host. An early study demonstrated that germ-free rabbits inoculated with a mouse gut

microbiota exhibited impaired gastrointestinal function compared to those given a normal rabbit microbiota [66]. Together, these functional studies and others suggest that interactions between hosts and their microbiota are not random and instead occur at various functional levels.

Here, we add an evolutionary component to these ideas by demonstrating that microbial communities from more evolutionarily distant hosts can be prone to more pronounced reductions in host performance or fitness. Specifically, *Peromyscus* deer mice inoculated with microbial communities from more distantly related species tended to exhibit lower food digestibility. The exception to this trend was the *P. eremicus* to *P. polionotus* group, which did not exhibit any decrease in digestibility. It should be noted that *P. eremicus* also did not follow phylosymbiosis (Fig 4B), which may explain the departure from our expected trend in digestibility. For example, deviations from phylosymbiosis could be due to a microbial community assembly that is inconsequential to host digestibility. Therefore, transferring a nonphylosymbiotic community between host species may not yield performance costs.

An alternative explanation for our results could be that hosts are acclimated to their established microbiota, and the introduction of foreign microbiota either elicits a host immune response or disrupts the established microbiota, thus decreasing digestibility. One technique to distinguish between adaptation and acclimation would be to conduct experiments in germ-free *P. polionotus* recipients. However, the derivation of germ-free mammals is a difficult and expensive process [67] and has not been conducted for *Peromyscus*. Earlier studies utilizing germ-free mammals demonstrate that microbial communities from evolutionarily distant hosts negatively impact gastrointestinal function [66] and immune development [68], thus supporting our hypothesis of functional matching between host and the gut microbiota.

Additionally, among very closely related species, *Nasonia* exposed to interspecific microbiota have lower fitness than those exposed to intraspecific microbiota. While this experiment utilized heat-killed bacteria to avoid shifts in the microbiota composition during media growth, the protocol is sufficient to test the predictions of phylosymbiosis. First, isolated microbial products can exert drastic effects on eukaryotic partners. For example, a sulfonolipid purified from bacteria can induce multicellularity in choanoflagellates [69]. Additionally, the insect immune system can respond with strain-level specificity to heat-killed bacteria [70]. Therefore, we hypothesize that each *Nasonia* host species evolved to the products of their own gut microbiota rather than those of gut microbiota from related host species. Together, results from the *Peromyscus* and *Nasonia* functional experiments reveal the importance of host evolutionary relationships when considering interactions between hosts and their gut microbial communities and ultimately the symbiotic processes that can drive adaptation and speciation [71,72]. The molecular mechanisms underlying the functional bases of phylosymbiosis in various systems demand further studies.

Overall, we have established phylosymbiosis as a common, though not universal, phenomenon under controlled rearing with functional effects on host performance and survival. It is worth emphasizing again that this term is explicit and different from many other similar terms, such as coevolution, cospeciation, cocladogenesis, or codiversification [73]. While cospeciation of hosts and specific environmentally or socially acquired microbes—e.g., hominids and gut bacterial species [74] or the bobtail squid and *Vibrio* luminescent bacteria [75]—could contribute in part to phylosymbiosis, concordant community structuring with the host phylogeny is not dependent on parallel gene phylogenies but instead on total microbiota compositional divergence. Phylosymbiosis does not assume congruent splitting from an ancestral species because it does not presume that microbial communities are stable or even vertically transmitted from generation to generation [1,12]. Rather, phylosymbiosis predicts that the congruent relationships of host evolution and microbial community similarities could have

varied assembly mechanisms in space and time and be newly assembled each generation (though see our discussion of transmission routes above). Moreover, the findings here imply that across wide-ranging evolutionary timescales and animal systems, there is a functional eco-evolutionary basis for phylosymbiosis, at least under controlled conditions.

It may be difficult to detect phylosymbiosis in natural populations because of extensive environmental variation that overwhelms the signal. We suggest that one way to potentially overcome this challenge is to start with laboratory-controlled studies that identify (i) phylosymbiotic communities and (ii) the discriminating microbial taxa between host species. Resultantly, investigations can test whether these microbial signatures exist in natural populations, albeit perhaps in a smaller fraction of the total microbiota that is mainly derived by environmental effects. Another advantage of controlled studies is that the functional effects, both positive and negative, of a phylosymbiotic community assembly can be carefully measured in the context of host evolutionary history.

Materials and Methods

Ethics Statement

Procedures involving functional microbiota transplants in *Peromyscus* mice were approved by the University of Utah Institutional Animal Care and Use Committee under protocol 12–12010. Mice obtained from the *Peromyscus* Genetic Stock Center were reared under IACUC approved protocols, and only fecal samples were directly utilized. While our paper contains data for several primate species, this data was conducted by another research group, has been previously published, and is now publicly available. Thus, there was no requirement of approved protocols for the primate species.

Nasonia Husbandry and Sample Collection

Nasonia were reared as previously described [2]. Four strains were used: *Nasonia vitripennis* (strain 13.2), *N. longicornis* (IV7U-1b), *N. giraulti* (RV2x(u)), *N. oneida* (NAS_NONY(u)). To collect individuals for microbiota analysis, virgin females were sorted as pupae into sterile glass vials and collected within the first 24 h of eclosing as adults. Subsequently, they were rinsed with 70% ETOH for 2 min, a 1:10 bleach solution for 2 min, followed by two rinses in sterile water. Individuals were then placed in 1.5 ml tubes and flash frozen in liquid nitrogen. They were then stored at -80°C until DNA extractions. Fifty individuals were collected per strain.

Drosophila Husbandry and Sample Collection

Nine strains of *Drosophila* were obtained from the University of California San Diego *Drosophila* Species Stock Center. Six strains were used in the microbiome analysis because they were *Wolbachia*-free: *Drosophila melanogaster* (Strain Dmel, stock number 14021–0248.25), *D. simulans* (Dsim, 14021–0251.195), *D. yakuba* (Dyak, 14021–0261.01), *D. erecta* (Dere, 14021–0224.01), *D. pseudoobscura* (Dpse, 14011–121.94), and *D. mojavensis* (Dmow, 15081–1352.22). The three strains that tested positive for *Wolbachia* (method described below) were: *D. sechellia* (14021–0248.25), *D. ananassae* (14021–0371.13), and *D. willistoni* (14030–0811.24). All strains were reared on a cornmeal media (*Drosophila* Species Stock Center: http://stockcenter.ucsd.edu/info/food_cornmeal.php) with a sterile Braided Dental Roll (No. 2, Crosstex, Atlanta, Georgia, US) inserted into the surface of the media. All stocks were incubated at 25°C with a 12-h light–dark cycle and monitored every 24 h. Every 14 d, stock vials were cleared of any emerged adults, and 6 h later, ten virgin females and three males were transferred to new food

vials. This conditioning on the same food was done for five generations before setting up media vials for sample collection. For each of the six strains, five virgin females were mated with two males and allowed to oviposit for 24 h; afterwards, the parents were removed and the vials were incubated as per above.

After 12 d, vials were cleared and virgin females were collected every 4–6 h over a 36-h period. All females were rinsed with 70% ETOH for 2 min, a 1:10 bleach solution for 2 min, followed by two rinses in sterile water. Individual adult flies were then placed in 1.5 ml tubes and flash frozen in liquid nitrogen. They were then stored at -80°C until DNA extractions. Approximately 25–30 virgin adult females were collected per strain.

Mosquito Husbandry and Sample Collection

Mosquitoes were acquired from the Malaria Research and Reference Reagent Resource Center as eggs on damp filter paper within 24 h of being laid. Eight strains were used: *Anopheles funestus* (strain name FUMOZ), *An. farauti* s.s. (FAR1), *An. quadrimaculatus* (GORO), *An. arabiensis* (SENN), *An. gambiae* (MALI NIH), *Aedes aegypti* (COSTA RICA), *Ae. albopictus* (ALBO), and *Culex tarsalis* (YOLO F13). Eggs were floated in 350 ml of sterile water with 1.5 ml of 2% yeast slurry and autoclaved within a sterile and lidded clear plastic container. Containers were enclosed within a larger sterile clear container and placed inside an incubator set at 25°C with a 12-h light–dark cycle and monitored every 24 h. After 48 h, the hatched larvae were sorted out and 100–150 of each species were placed in new sterile water (150 ml) with 30 mg of powdered koi food (Laguna Goldfish & Koi all season pellets). Water level was maintained at 150 ml, and larvae were fed 30 mg of powdered koi food every day for a total of 13 d. All pupae were discarded (frozen and autoclaved) on day 10, and new pupae were collected every 12 h on day 11, 12, and 13. Water samples were also collected and frozen for microbial analysis on day 11.

To collect individuals for microbiota analysis, pupae were sorted according to sex, and all females were rinsed with 70% ETOH for two min, then 1:10 bleach solution for two min, followed by two rinses in sterile water. Individual pupae were then placed in 1.5 ml tubes and flash frozen in liquid nitrogen. They were then stored along with their corresponding water sample at -80°C until DNA extractions. Ten to 25 individuals were collected per strain.

Peromyscus Husbandry and Sample Collection

Fecal samples were collected from the *Peromyscus* Genetic Stock Center at the University of South Carolina. Six stock species of *Peromyscus* were used: *P. maniculatus* (stock BW), *P. polionotus subgriseus* (PO), *P. leucopus* (LL), *P. californicus insignis* (IS), *P. aztecus hylocetes* (AM), and *P. eremicus* (EP). All mice were reared using their standard care practices at the stock center on the same mouse chow diet. Cages were cleaned at regular intervals for all species, and all species were caged within the same facility. Individuals from nonmating cages of females (five to six per cage) were used for collections.

Fecal pellets were collected on a single morning from individual mice directly into a sterile tube and placed on dry ice before being stored at -80°C for 24 h. Samples were then shipped overnight on dry ice and again stored at -80°C until DNA extractions. One to three pellets from 15 individuals were collected per strain.

In order to eliminate the introduction of confounding factors and exclude any subjects that had a pinworm infection at the time of sample collection, we conducted a screen to confirm the pinworm status of each mouse. Pinworm status was confirmed by PCR. Primers utilized to amplify the 28S rDNA D1 and D2 domains of multiple pinworm species were developed and confirmed with positive DNA samples of *Syphacia obvelata* and *Aspiculuris tetraptera* (received

from the Feldman Center for Comparative Medicine at the University of Virginia). The C1 primer 5'-ACCCGCTGAATTTAAGCAT-3' and the D1 primer 5'-TCCGTGTTTCAA-GACGG-3' were amplified under the following reaction conditions: 94°C for 1 min; 35 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 30 s; and a final elongation time at 72°C for 2 min. The resultant samples were then visualized on a 1% agarose gel. Of the 84 fecal specimens analyzed, 8 of the samples showed amplification at 750 bp corresponding to the expected amplification size of the pinworm DNA sequence. For confirmation, the 750 bp bands were extracted using a Wizard Gel Extraction Kit (Promega Corporation, Madison, Wisconsin, US) and sequenced (GENEWIZ, Inc, New Jersey, US). Sequence results confirmed the presence of *Aspiculuris tetraptera* infection, and these 8 samples and were excluded from further analysis.

Wolbachia Screens of Stock Insect Lines

The presence or absence of *Wolbachia* was checked using two replicates of three individuals per species. DNA extraction was performed with PureGene DNA Extraction Kit (Qiagen), and fragments of the 16S rDNA gene were PCR amplified using primer set WolbF and WolbR3 [76]. Only stock strains that were *Wolbachia* negative were used in the experiments.

Insect DNA Extraction

Individual insects (and the mosquitoes' corresponding water samples) were mechanically homogenized with sterile pestles while frozen within their collection tube. The samples were then thawed to room temperature for 30 s and flash frozen again in liquid nitrogen with additional mechanical homogenization. The samples were finally processed using the ZR-Duet DNA/RNA MiniPrep Kit (Zymo Research, Irvine, California, US). Samples were then quantified using the dsDNA BR Assay kit on the Qubit 2.0 Fluorometer (Life Technologies).

DNA Isolation from Mouse Samples

The PowerSoil DNA isolation kit (Mo Bio Laboratories, Carlsbad, California, US), was utilized to extract DNA from 20 mg of mouse fecal material per sample according to manufacturer's protocol after being mechanically homogenized with sterile pestles while frozen within their collection tube. Samples were then quantified using the dsDNA BR Assay kit on the Qubit 2.0 Fluorometer.

PCR, Library Prep, and Sequencing

Total genomic DNA was quantified using dsDNA HS Assay kit on the Qubit. Using two μ l of DNA, a 20 μ l PCR reaction of 28S general eukaryotic amplification was conducted on each sample, with only 25 cycles. Products were purified using Agencourt AMPure XP, quantified using the dsDNA HS Assay kit on the Qubit, and compared to the amount of 16S amplification from the same DNA volume and PCR reaction volume as previously described [2]. PCR amplification of the bacteria 16S rRNA was performed with the 27F 5'-AGAGTTTGATCCTGGCT-CAG-3' and 338R 5'-GCTGCCTCCCGTAGGAGT-3' "universal" bacterial primers with the NEBNext High-Fidelity 2X PCR Master Mix; duplicate reactions were generated per sample, which were pooled together postamplification. For sequencing runs 1 (*Peromyscus*) and 2 (*Nasonia*, mosquito, and *Drosophila*), 16S PCR products that were made into libraries had their concentrations normalized relative to about 1,000 ng/ml and 2,000 ng/ml of the 28S quantity for library prep respectively.

Using the Encore 384 Multiplex System (NuGEN, San Carlos, California, US), each samples' 16S product was ligated with Illumina NGS adaptors and a unique barcode index (after

the reverse adaptor). The samples were then purified using Agencourt AMPure XP and quantified using the dsDNA HS Assay kit on the Qubit. Samples were subsequently pooled.

Each pooled library was run on the Illumina MiSeq using either the MiSeq Reagent Kit V2 or V3 for paired-end reads. Run 1 was conducted at the University of Georgia Genomics Facility and run 2 was conducted at Vanderbilt Technologies for Advanced Genomics (VANTAGE).

Sequence Quality Control

Sequence quality control and OTU analyses were carried out using QIIME version 1.8.0 [77]. Forward and reverse paired-end sequences were joined and filtered if they met the following criteria: they fell below an average Phred quality score of 25, contained homopolymer runs or ambiguous bases in excess of 6 nucleotides, or were shorter than 200 base pairs. Sequences were also removed if there were errors in the primer sequence or if barcodes contained errors and could not be assigned to a sample properly. A total of 5,065,121 reads passed quality control for the meta-analysis, with an average read length of 310 ± 48 nucleotides. *Drosophila*: 648,676 reads, average length 315 ± 23 . hominid: 1,292,542 reads, average length 247 ± 38 . mosquito: 664,350 reads, average length 328 ± 19 . *Nasonia*: 864,969 reads, average length 322 ± 15 . *Peromyscus*: 295,752 reads, average length 347 ± 12 .

OTU Analysis

Chimeric sequences were evaluated and removed using the UCHIME algorithm [78] for the intersection of de novo and GreenGenes 13_5 non-chimeras [79]. The sequences were then clustered into OTUs at 94%, 97%, and 99% similarity using the USEARCH open-reference method [80]. OTUs were mapped at the respective percent against the GreenGenes 13_5 database and screened for a minimum group size of two counts, with dereplication based on full sequences [79]. Representative sequences were chosen as the most abundant representative in each OTU cluster and aligned using GramAlign [81]. A phylogenetic tree of the representative sequences was built in QIIME [77] with the FastTree method and midpoint rooting [82]. Taxonomy was then assigned to the OTU representatives with the UCLUST method against the GreenGenes 13_5 database [79]. OTU tables were constructed in QIIME [77] and sorted by sample IDs alphabetically.

Sample and OTU Quality Control

OTU tables were screened to remove any OTUs classified as chloroplast, unassigned, and *Wolbachia*. Individual samples were assessed for low sequence coverage affecting community profiles and diversity as well as for processing errors based on minimum count thresholds assessed against group means. Following rarefaction, counts were subsequently chosen as the highest rarefaction number allowed by the smallest sample's count representation in each respective clade and the meta-analysis. Alpha diversity was measured using Shannon and Chao1 metrics generated with the QIIME alpha_rarefaction script. Plots of alpha diversity at a range of rarefied levels were used to assess and remove samples with low diversity.

Meta-Analysis

The PCoA (Fig 2A) components for the meta-analysis were constructed using the QIIME jackknifed_beta_diversity script. The OTU table first underwent rarefaction, followed by the computation of Bray-Curtis beta diversity distances for each rarefied table. PCoA plots of the first three coordinate dimensions were generated using a custom Python script. Individual samples are each depicted as a point and are colored by host clade of origin.

The community profile (Fig 2B) for the meta-analysis was generated using a custom Python script and BIOM tools [83]. OTU tables were first converted to relative abundance for each sample, and bacterial taxonomy was collapsed at the class level. Bacterial classes were sorted alphabetically, and a stacked bar chart representing the relative abundance for each sample was constructed.

The network analysis (Fig 2C) was visualized using Cytoscape [84]. OTU tables were first collapsed by bacterial taxonomy at the genus level, and QIIME's `make_otu_network` script was used to construct connections between each bacterial genus to individual hosts based on relative abundance. Network files were then imported into Cytoscape, where the network was computed using an edge-weighted force directed layout. Nodes were colored by host clade, and connections were colored by key bacterial phylum observed in high abundance (i.e., Actinobacteria, Bacteroidetes, Firmicutes, Proteobacteria) and gray for additional phylum.

Alpha diversity plots (Fig 2D) were prepared using the Phyloseq package [85]. OTU tables collapsed by host species were imported into Phyloseq, and the `plot_richness` function was used to generate box-and-whisker plots of Shannon alpha-diversity. Plots were colored by host clade of origin.

Microbiota Dendrograms

Microbiota dendrograms were constructed using the QIIME `jackknifed_beta_diversity` script. OTU table counts were first collapsed by host species of origin to get representative species microbiota profiles. The pipeline script performed 1,000 rarefactions on each table and calculated Bray-Curtis beta diversity distances for each. Bray-Curtis distance matrices were UPGMA clustered to give dendrograms of interspecific relatedness. The role of 97% versus 99% OTU clustering cutoffs and weighted and unweighted UniFrac beta diversity measures (S2 Fig) were evaluated for Robinson-Foulds and matching cluster congruence with host phylogeny.

Host Phylogenies

Host phylogenetic trees were constructed using sequences for each host species' cytochrome oxidase gene downloaded from the NCBI. COI was chosen as a highly conserved molecular marker, and it is widely used for interspecific phylogenetic comparison [86]. Sequences were initially aligned using Muscle v3.8.31 [87]. Gap positions generated through inserts and deletions were removed, and overhanging sequence on 5' and 3' ends were trimmed. Models of molecular evolution were evaluated using jModelTest v2.1.7 [88], and the optimal model was used for final alignment and tree building in RaxML v8.0.0 [89]. The *Nasonia* and *Peromyscus* clades were carried out using the same methodology—except for final alignment and tree building in PhyML v3.0 [90]—and for *Peromyscus* the AVPR1A gene was concatenated with COI to further resolve the phylogeny. All trees are concordant with well-established phylogenies from literature references noted in the Results section.

Robinson-Foulds and Matching Cluster Congruency Analyses

Quantifying congruence between host phylogeny and microbiota dendrogram relationships (Fig 4) was carried out with a custom Python script and the TreeCmp program [91]. The topologies of both trees were constructed, and the normalized Robinson-Foulds score [35] and normalized matching cluster score [34] were calculated as the number of differences between the two topologies divided by the total possible congruency score for the two trees. Next, 100,000 random trees were constructed with the same number of leaf nodes, and each was compared to the host phylogeny. The number of trees which had an equivalent or

better score than the actual microbiota dendrogram were used to calculate the significance of observing that topology under stochastic assembly. Normalized results of both statistics have been provided to facilitate comparison. Matching cluster and Robinson–Foulds p -values were determined by the probability of 100,000 randomized bifurcating dendrogram topologies yielding equivalent or more congruent phylosymbiotic patterns than the microbiota dendrogram.

Intraspecific Versus Interspecific Beta Diversity Distances

Within each clade, the Bray–Curtis distances calculated by the `jackknife_beta_diversity` script (Fig 3A) were separated by those that compared microbiota within a host species and those that compared between host species. The box-and-whisker plots were constructed in Python. Coloring indicates host clade of origin, and all intraspecific and interspecific distances are represented for each clade. These distances were then compared between the groups using a non-parametric, two-tailed Mann–Whitney U test implemented in SciPy [92,93].

ANOSIM Clustering

To evaluate intraspecific clustering (Fig 3B), the ANOSIM test was used to calculate the distinguishability of Bray–Curtis distances based on species of origin. Bray–Curtis distance matrices were generated using the QIIME `jackknifed_beta_diversity` script on tables of individuals rarefied 1,000 times. The QIIME script `compare_categories` was used to calculate ANOSIM scores using the Bray–Curtis distance matrix and host species as categories. 1,000 permutations were used to calculate the significance of clustering for each clade. Three-dimensional PCoA plots were generated in Python using components generated from Bray–Curtis distance matrices in QIIME, and the first three components are shown. Points are colored by host species within each clade, and colors correlate with the species labels in Fig 4 for reference.

Correlation of ANOSIM Clustering and Clade Age

A general linear regression was performed to test the correlation between age of clade origin and the intraspecific clustering measured through ANOSIM R-statistic scores. Cladogenesis Age was Log10 transformed to normalize the distance scale between samples (1, 10, 100 MYA). The regression was carried out in Stata v12.0 to determine the coefficient (R^2) and significance (p -value).

Random Forest Analyses

OTU tables were first collapsed at each bacterial taxonomic level (i.e., phylum . . . genus) using the QIIME script `summarize_taxa`. Then, both the raw OTU table and each collapsed table underwent ten rarefactions to an even depth using the QIIME script `multiple_rarefactions_even_depth`. RFC models were constructed with the `supervised_learning` script for 1,000 rounds of ten-fold Monte Carlo cross validation on each table. At each level, the results were collated and averages were taken for the ten rarefied tables. Host species were used as the category for RFC model distinguishability, testing the ability to assign samples to their respective host species. The average class error for each clade was subtracted from 100 to get the percent accuracy of the models at each taxonomic level. The same methodology was used for constructing RFC models for the meta-analysis, with the only exception being that host species, host clade, and vertebrate or invertebrate categories were tested for distinguishability.

Microbiota Transplants

Peromyscus. We tested the effects of allochthonous microbial communities on host performance by conducting a series of microbial transplants from various donor rodent species into a single recipient species, the oldfield mouse (*Peromyscus polionotus*). We obtained virgin, female *Peromyscus* species (*P. polionotus*, *P. maniculatus*, *P. leucopus*, *P. eremicus*, *P. californicus*) from the *Peromyscus* stock center. We also obtained three female individuals of *Neotoma lepida* (*Neotoma* is the sister genus of *Peromyscus*) from Dr. M. Denise Dearing (University of Utah). Additionally, we obtained six female individuals of outbred *Mus musculus* from Dr. Wayne Potts (University of Utah). The founding animals of this colony were collected from near Gainesville, Florida, US, and the animals have been randomly bred in captivity for roughly 13 generations and are still highly outbred [94,95]. All rodent species were maintained on powdered laboratory rodent chow (Formula 8904, Harlan Teklad, Madison, Wisconsin, US) except for woodrats, which were fed powdered rabbit chow (Formula 2031, Harlan Teklad, Madison, Wisconsin, US), given that woodrats are herbivorous. All procedures involving rodents were approved under the University of Utah Institutional Animal Care and Use Committee protocol #12–12010.

To conduct microbial transplants, we followed a protocol that was previously established to transplant the microbiota from *Neotoma lepida* into *Rattus norvegicus* [39]. First, donor feces were collected from three to six individuals of each donor species by placing rodents in wire-bottom metabolic cages overnight and collecting feces the next morning. Feces were then ground with a mortar and pestle and mixed into powdered laboratory chow (Formula 8904, Harlan Teklad, Madison, Wisconsin, US) at a ratio of 15% w/w. Recipient animals (five to six individuals per group) were fed food containing feces of a particular donor species for two nights. Then, recipient animals were fed normal laboratory diets for 6 d, which is a sufficient time for the clearance of transient, ingested microbes [96]. We then measured food intake and dry matter digestibility by placing animals into wire-bottom metabolic cages. Animals were presented with a known amount of powdered rodent chow overnight. The next morning, remaining food was weighed, and feces were collected, dried overnight, and weighed. Food intake was calculated as g dry food presented–g dry food remaining. Dry matter digestibility was calculated as (g dry food ingested–g dry feces produced) / g dry food ingested.

We investigated whether microbial communities from more distantly related hosts affected performance metrics in recipients. We compared food intake using ANOVA and Tukey's HSD test across recipient groups. We also conducted correlations of dry matter digestibility and estimated divergence times based off of previously published phylogenies [37,38]. We performed correlations using both untransformed divergence times and log-transformed divergence times.

Nasonia. We tested the effects of allochthonous microbial communities on host survival by exposing two recipient species (*N. vitripennis* or *N. giraulti*) to a suspension of heat-killed microbes isolated from three donor *Nasonia* species (*N. vitripennis*, *N. giraulti*, and *N. longicornis*). We reared *Nasonia* in an in vitro rearing system [41] and inoculated germ-free larvae in 6 mm diameter transwell inserts with autochthonous microbiota, allochthonous microbiota, and sterile phosphate-buffered saline (PBS) for the first 8 d after embryo hatching. Microbiota were purified from fourth instar larvae of *Nasonia* by filtration through a 5 μ m filter and centrifugation at 10,000 rpm for 3 min. The pellet was suspended in a sterile PBS solution at a concentration of 5×10^6 CFU of microbiota bacteria (determined by tryptic soy agar plating) per milliliter. 20 μ L of this microbiota suspension was added to the transwell inserts for each of the 8 inoculation days. *Nasonia* rearing media was replaced daily just before the inoculations.

Measurements of *Nasonia* survival from first instar larvae to adulthood were determined using transwell insert images taken with an AmScope MT1000 camera. For each transwell, live

larval counts were recorded 3 d post-embryo hatching. Adult counts were determined by recording the number of remaining larvae and pupae in each transwell sample 20 d after embryo hatching (5–7 d after first adult eclosion) and subtracting that number from the larval counts previously recorded. Normalized adult survival per transwell sample was calculated as the percent survival of *Nasonia* from 3 d to 20 d after embryo hatching divided by the average percent survival of the autochthonous microbiota treatment group. We compared survival between the autochthonous and allochthonous treatment groups using Mann–Whitney U tests.

Supporting Information

S1 Fig. Comparisons of intraspecific and interspecific Bray-Curtis distances for pairwise combinations of all species. Bray-Curtis beta diversity distances were computed for all pairs of individuals within each clade from 99 percent OTUs. Colored circles denote the named species, and colors within box-and-whisker plots denote to which species it is being compared. Boxes represent the 25th to 75th quartiles with the central line depicting the group median, and whiskers showing the 1.5 interquartile extent. Data available at [96] in folder Fig_3A_&_S1. (PDF)

S2 Fig. Phylosymbiosis analysis for alternative beta-diversity metrics and OTU clustering cutoffs. The normalized Robinson-Foulds metric and the normalized Matching Cluster metric were used to evaluate the congruence between host phylogenies and microbiota dendrograms for Bray Curtis, Unweighted UniFrac, and Weighted UniFrac beta-diversity metrics at both 97 and 99 percent clustered OTUs. Data available at [96] in folder Fig_S2. (PDF)

S3 Fig. Fine-resolution overlap between donor and recipient microbial communities. White bars represent shared OTUs between donor and recipients and thus the possible range of transfer. Colored bars represent the portion of shared OTUs that are donor-specific and thus transfer of unique OTUs between donor and recipients. Panels (A) and (B) depict the mean \pm s.e.m. percentage of OTUs. Panels (C) and (D) show the mean \pm s.e.m. abundance of total sequences. These analyses were conducted with OTU-picking at both 97% and 99% sequence identities. Data available at [96] in folder Fig_S3. (TIFF)

S4 Fig. Effects of allochthonous versus autochthonous microbial communities on the food intake of recipient mice. Divergence times between *P. polionotus* and donor species were determined from previously published phylogenies [38, 39]. Points represent mean values \pm s. e.m. for each group (n = 5–6 recipients per group). Data available at [96] in folder Fig_5_&_S4. (TIFF)

S1 Table. Table of Random Forest accuracy in classifying the microbiota by host species in each host clade, and by host species, clade, and mammal or invertebrate taxonomy in the meta-analysis. Models were generated using OTUs or abundance collapsed by bacterial taxonomy. Red boxes highlight the highest classification accuracy. Ten-fold cross validation assessed the percent classification accuracy for test sets excluded from model training. (XLSX)

S2 Table. Table of Random Forest model mean decrease in accuracy when genera are excluded from classification of the microbiota in each host clade. Random Forest models were generated using genera collapsed bacterial taxonomies. Genera are ordered by those that

contribute the most accuracy to the model to those that contribute the least accuracy to the model, measured in the form of decrease in model accuracy when a genus is excluded from model construction. Standard deviations of mean decrease in model accuracy are also provided.

(XLSX)

S3 Table. Tables of microbiota taxon in the meta-analysis with varying abundance between host clades or between vertebrates and invertebrates. The meta-analysis OTU table was collapsed at each bacterial taxonomic level (Phylum to Genus), and converted to relative abundance. Kruskal-Wallis tests were performed on microbial taxon within each table, testing for differences in the mean abundance across host clades or vertebrates and invertebrates. The results were sorted from high to low significance of p-values, which are provided alongside False Discovery Rate and Bonferroni corrected p-values. Mean abundances of each taxon within host clades or vertebrates and invertebrates are provided as a heatmap, with dark blue indicating high abundance, light blue centered at the 5% most abundant values and fading to white for low abundance or non-existent taxon.

(XLSX)

Acknowledgments

We thank MR4 for providing mosquito eggs (contributed by Sandra Allan, Nora Besansky, Mustafa Dakeen, William Collins, Maureen Coetzee, William Reisen, William Brogdon, and MR4 Vector Activity). We would like to especially recognize the substantial efforts made by Michael Felder and Janet Crossland for evaluating the stock lines and collecting fecal samples from the *Peromyscus* Genetic Stock Center at the University of South Carolina, Dr. Sanford H. Feldman of the Center for Comparative Medicine at the University of Virginia for providing the pinworm DNA, and Theresa Bark for screening the *Peromyscus* stool for pinworm DNA. We thank Dr. M. Denise Dearing for supplying equipment and facilities to conduct the *Peromyscus* functional experiments and Dr. Hopi Hoekstra for help with the *Peromyscus* host phylogeny. We would like to thank Ran Blekhman and Michael Burns for support in processing preliminary data and Sarah Bordenstein, Bojana Jovanovic, and Lisa Funkhouser for providing feedback on an earlier version of the manuscript.

Author Contributions

Conceptualization: Andrew W. Brooks, Kevin D. Kohl, Robert M. Brucker, Edward J. van Opstal, Seth R. Bordenstein.

Data curation: Andrew W. Brooks, Robert M. Brucker.

Formal analysis: Andrew W. Brooks, Kevin D. Kohl, Robert M. Brucker, Edward J. van Opstal.

Funding acquisition: Kevin D. Kohl, Robert M. Brucker, Seth R. Bordenstein.

Investigation: Andrew W. Brooks, Kevin D. Kohl, Robert M. Brucker, Edward J. van Opstal, Seth R. Bordenstein.

Methodology: Andrew W. Brooks, Kevin D. Kohl, Robert M. Brucker, Edward J. van Opstal, Seth R. Bordenstein.

Project administration: Robert M. Brucker, Seth R. Bordenstein.

Resources: Robert M. Brucker, Seth R. Bordenstein.

Software: Andrew W. Brooks, Robert M. Brucker, Seth R. Bordenstein.

Supervision: Robert M. Brucker, Seth R. Bordenstein.

Validation: Andrew W. Brooks, Robert M. Brucker, Edward J. van Opstal, Seth R. Bordenstein.

Visualization: Andrew W. Brooks, Kevin D. Kohl, Robert M. Brucker, Edward J. van Opstal, Seth R. Bordenstein.

Writing – original draft: Andrew W. Brooks, Kevin D. Kohl, Robert M. Brucker.

Writing – review & editing: Andrew W. Brooks, Kevin D. Kohl, Robert M. Brucker, Edward J. van Opstal, Seth R. Bordenstein.

References

1. Brucker RM, Bordenstein SR. The roles of host evolutionary relationships (genus: *Nasonia*) and development in structuring microbial communities. *Evol.* 2012; 66:349–62.
2. Brucker RM, Bordenstein SR. The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus *Nasonia*. *Science.* 2013; 341:667–9. doi: [10.1126/science.1240659](https://doi.org/10.1126/science.1240659) PMID: [23868918](https://pubmed.ncbi.nlm.nih.gov/23868918/)
3. Burns AR, Stephens WZ, Stagaman K, Wong S, Rawls JF, Guillemin K, et al. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME Journal.* 2015; 10:655–64. doi: [10.1038/ismej.2015.142](https://doi.org/10.1038/ismej.2015.142) PMID: [26296066](https://pubmed.ncbi.nlm.nih.gov/26296066/)
4. David LA, Materna AC, Friedman J, Campos-Baptista MI, Blackburn MC, Perrotta A, et al. Host lifestyle affects human microbiota on daily timescales. *Genome Biol.* 2014; 15:R89. doi: [10.1186/gb-2014-15-7-r89](https://doi.org/10.1186/gb-2014-15-7-r89) PMID: [25146375](https://pubmed.ncbi.nlm.nih.gov/25146375/)
5. Franzenburg S, Walter J, Künzel S, Wang J, Baines JF, Bosch TC, et al. Distinct antimicrobial peptide expression determines host species-specific bacterial associations. *Proc Natl Acad Sci.* 2013; 110: E3730–8. doi: [10.1073/pnas.1304960110](https://doi.org/10.1073/pnas.1304960110) PMID: [24003149](https://pubmed.ncbi.nlm.nih.gov/24003149/)
6. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. Evolution of mammals and their gut microbes. *Science.* 2008; 320:1647–51. doi: [10.1126/science.1155725](https://doi.org/10.1126/science.1155725) PMID: [18497261](https://pubmed.ncbi.nlm.nih.gov/18497261/)
7. Muegge BD, Kuczynski J, Knights D, Clemente JC, González A, Fontana L, et al. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science.* 2011; 332:970–4. doi: [10.1126/science.1198719](https://doi.org/10.1126/science.1198719) PMID: [21596990](https://pubmed.ncbi.nlm.nih.gov/21596990/)
8. Org E, Parks BW, Joo JW, Emert B, Schwartzman W, Kang EY, et al. Genetic and environmental control of host-gut microbiota interactions. *Genome Res.* 2015; 25:1558–69. doi: [10.1101/gr.194118.115](https://doi.org/10.1101/gr.194118.115) PMID: [26260972](https://pubmed.ncbi.nlm.nih.gov/26260972/)
9. Rawls JF, Mahowald MA, Ley RE, Gordon JI. Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell.* 2006; 127:423–33. doi: [10.1016/j.cell.2006.08.043](https://doi.org/10.1016/j.cell.2006.08.043) PMID: [17055441](https://pubmed.ncbi.nlm.nih.gov/17055441/)
10. Sanders JG, Powell S, Kronauer DJC, Vasconcelos HL, Frederickson ME, Pierce NE. Stability and phylogenetic correlation in gut microbiota: lessons from ants and apes. *Mol Ecol.* 2014; 23:1268–83. doi: [10.1111/mec.12611](https://doi.org/10.1111/mec.12611) PMID: [24304129](https://pubmed.ncbi.nlm.nih.gov/24304129/)
11. Davenport ER. Elucidating the role of the host genome in shaping microbiome composition. *Gut Microbes.* 2016; 7:178–84. doi: [10.1080/19490976.2016.1155022](https://doi.org/10.1080/19490976.2016.1155022) PMID: [26939746](https://pubmed.ncbi.nlm.nih.gov/26939746/)
12. Brucker RM, Bordenstein SR. Speciation by symbiosis. *Trends Ecol Evol.* 2012; 27:443–51. doi: [10.1016/j.tree.2012.03.011](https://doi.org/10.1016/j.tree.2012.03.011) PMID: [22541872](https://pubmed.ncbi.nlm.nih.gov/22541872/)
13. Easson CG, Thacker RW. Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Front Microbiol.* 2014; 5:532. doi: [10.3389/fmicb.2014.00532](https://doi.org/10.3389/fmicb.2014.00532) PMID: [25368606](https://pubmed.ncbi.nlm.nih.gov/25368606/)
14. Phillips CD, Phelan G, Dowd SE, McDonough MM, Ferguson AW, Hanson JD, et al. Microbiome analysis among bats describes influences of host phylogeny, life history, physiology, and geography. *Mol Ecol.* 2012; 21:2617–27. doi: [10.1111/j.1365-294X.2012.05568.x](https://doi.org/10.1111/j.1365-294X.2012.05568.x) PMID: [22519571](https://pubmed.ncbi.nlm.nih.gov/22519571/)
15. Moeller AH, Li Y, Ngole EM, Ahuka-Mundede S, Lonsdorf EV, Pusey AE, et al. Rapid changes in the gut microbiome during human evolution. *Proc Natl Acad Sci.* 2014; 111:16431–5. doi: [10.1073/pnas.1419136111](https://doi.org/10.1073/pnas.1419136111) PMID: [25368157](https://pubmed.ncbi.nlm.nih.gov/25368157/)
16. Ochman H, Worobey M, Kuo CH, Ndjanga JBN, Peeters M, Hahn BH, et al. Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biol.* 2010; 8:e1000546. doi: [10.1371/journal.pbio.1000546](https://doi.org/10.1371/journal.pbio.1000546) PMID: [21103409](https://pubmed.ncbi.nlm.nih.gov/21103409/)

17. Dietrich C, Köhler T, Brune A. The cockroach origin of the termite gut microbiota: patterns in bacterial community structure reflect major evolutionary events. *Appl Environ Microbiol.* 2014; 80:2261–9. doi: [10.1128/AEM.04206-13](https://doi.org/10.1128/AEM.04206-13) PMID: [24487532](https://pubmed.ncbi.nlm.nih.gov/24487532/)
18. Chandler JA, Lang JM, Bhatnagar S, Eisen JA, Kopp A. Bacterial communities of diverse *Drosophila* species: ecological context of a host-microbe model system. *PLoS Genet.* 2011; 7:e1002272. doi: [10.1371/journal.pgen.1002272](https://doi.org/10.1371/journal.pgen.1002272) PMID: [21966276](https://pubmed.ncbi.nlm.nih.gov/21966276/)
19. Wong ACN, Chaston JM, Douglas AE. The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *ISME Journal.* 2013; 7:1922–32. doi: [10.1038/ismej.2013.86](https://doi.org/10.1038/ismej.2013.86) PMID: [23719154](https://pubmed.ncbi.nlm.nih.gov/23719154/)
20. Staubach F, Baines JF, Künzel S, Bik EM, Petrov DA. Host species and environmental effects on bacterial communities associated with *Drosophila* in the laboratory and in the natural environment. *PLoS ONE.* 2013; 8:e70749. doi: [10.1371/journal.pone.0070749](https://doi.org/10.1371/journal.pone.0070749) PMID: [23967097](https://pubmed.ncbi.nlm.nih.gov/23967097/)
21. Hird SM, Sánchez C, Carstens BC, Brumfield RT. Comparative gut microbiota of 59 Neotropical bird species. *Front Microbiol.* 2015; 6:1403. doi: [10.3389/fmicb.2015.01403](https://doi.org/10.3389/fmicb.2015.01403) PMID: [26733954](https://pubmed.ncbi.nlm.nih.gov/26733954/)
22. Baxter NT, Wan JJ, Schubert AM, Jenior ML, Myers P, Schloss PD. Intra- and interindividual variations mask interspecies variations in the microbiota of sympatric *Peromyscus* populations. *Appl Environ Microbiol.* 2015; 81:396–404. doi: [10.1128/AEM.02303-14](https://doi.org/10.1128/AEM.02303-14) PMID: [25362056](https://pubmed.ncbi.nlm.nih.gov/25362056/)
23. Carmody RN, Gerber GK, Luevano JM, Gatti DM, Simes L, Svanson KL, et al. Diet dominates host genotype in shaping the murine gut microbiota. *Cell Host & Microbe.* 2015; 17:72–84.
24. Fraune S, Bosch TCG. Long-term maintenance of species-specific bacterial microbiota in the basal metazoan *Hydra*. *Proc Natl Acad Sci.* 2007; 104:13146–51. doi: [10.1073/pnas.0703375104](https://doi.org/10.1073/pnas.0703375104) PMID: [17664430](https://pubmed.ncbi.nlm.nih.gov/17664430/)
25. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J.* 1948; 27:379–423.
26. Brooks AW. Data Availability: Dryad Digital Repository. GitHub. 2016. <https://github.com/awbrooks19/phylosymbiosis>.
27. Neafsey DE, Waterhouse RM, Abai MR, Aganezov SS, Alekseyev MA, Allen JE, et al. Highly evolvable malaria vectors: the genomes of 16 *Anopheles* mosquitoes. *Science.* 2015; 347:1258522. doi: [10.1126/science.1258522](https://doi.org/10.1126/science.1258522) PMID: [25554792](https://pubmed.ncbi.nlm.nih.gov/25554792/)
28. Werren JH, Richards S, Desjardins CA, Niehuis O, Gadau J, Colbourne JK, et al. Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science.* 2010; 327:343–8. doi: [10.1126/science.1178028](https://doi.org/10.1126/science.1178028) PMID: [20075255](https://pubmed.ncbi.nlm.nih.gov/20075255/)
29. Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, et al. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature.* 2007; 450:203–18. doi: [10.1038/nature06341](https://doi.org/10.1038/nature06341) PMID: [17994087](https://pubmed.ncbi.nlm.nih.gov/17994087/)
30. Steiper ME, Young NM. Primate molecular divergence dates. *Mol Phylogenet Evol.* 2006; 41:384–94. doi: [10.1016/j.ympev.2006.05.021](https://doi.org/10.1016/j.ympev.2006.05.021) PMID: [16815047](https://pubmed.ncbi.nlm.nih.gov/16815047/)
31. Weber JN, Hoekstra HE. The evolution of burrowing behaviour in deer mice (genus *Peromyscus*). *Anim Behav.* 2009; 77:603–9.
32. Shin SC, Kim S-H, You H, Kim B, Kim AC, Lee K-A, et al. *Drosophila* microbiome modulates host developmental and metabolic homeostasis via insulin signaling. *Science.* 2011; 334:670–4. doi: [10.1126/science.1212782](https://doi.org/10.1126/science.1212782) PMID: [22053049](https://pubmed.ncbi.nlm.nih.gov/22053049/)
33. Everard A, Lazarevic V, Gaïa N, Johansson M, Ståhlman M, Backhed F, et al. Microbiome of prebiotic-treated mice reveals novel targets involved in host response during obesity. *ISME Journal.* 2014; 8:2116–30. doi: [10.1038/ismej.2014.45](https://doi.org/10.1038/ismej.2014.45) PMID: [24694712](https://pubmed.ncbi.nlm.nih.gov/24694712/)
34. Bogdanowicz D, Giaro K. On a matching distance between rooted phylogenetic trees. *Int J Appl Math Comp Sci.* 2013; 23:669–84.
35. Robinson DF, Foulds LR. Comparison of phylogenetic trees. *Math Biosci.* 1981; 53:131–47.
36. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol.* 2005; 71:8228–35. doi: [10.1128/AEM.71.12.8228-8235.2005](https://doi.org/10.1128/AEM.71.12.8228-8235.2005) PMID: [16332807](https://pubmed.ncbi.nlm.nih.gov/16332807/)
37. Platt RN, Amman BR, Keith MS, Thompson CW, Bradley RD. What Is *Peromyscus*? Evidence from nuclear and mitochondrial DNA sequences suggests the need for a new classification. *J Mamm.* 2015; 96:708–19.
38. Stepphan SJ, Adkins RM, Anderson J. Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Syst Biol.* 2004; 53:533–53. doi: [10.1080/10635150490468701](https://doi.org/10.1080/10635150490468701) PMID: [15371245](https://pubmed.ncbi.nlm.nih.gov/15371245/)
39. Kohl KD, Stengel A, Dearing MD. Inoculation of tannin-degrading bacteria into novel hosts increases performance on tannin-rich diets. *Environ Microbiol.* 2016; 18(6):1720–9. doi: [10.1111/1462-2920.12841](https://doi.org/10.1111/1462-2920.12841) PMID: [25753857](https://pubmed.ncbi.nlm.nih.gov/25753857/)

40. Munger JC, Karasov WH. Sublethal parasites and host energy budgets: tapeworm infection in white-footed mice. *Ecology*. 1989; 70:904–21.
41. Shropshire JD, Opstal EJ, Bordenstein SR. An optimized approach to germ-free rearing in the jewel wasp *Nasonia*. *PeerJ*. 2016; 4:e2316. doi: [10.7717/peerj.2316](https://doi.org/10.7717/peerj.2316) PMID: [27602283](https://pubmed.ncbi.nlm.nih.gov/27602283/)
42. Yun JH, Roh SW, Whon TW, Jung MJ, Kim MS, Park DS, et al. Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Appl Environ Microbiol*. 2014; 80:5254–64. doi: [10.1128/AEM.01226-14](https://doi.org/10.1128/AEM.01226-14) PMID: [24928884](https://pubmed.ncbi.nlm.nih.gov/24928884/)
43. Blekhman R, Goodrich JK, Huang K, Sun Q, Bukowski R, Bell JT, et al. Host genetic variation impacts microbiome composition across human body sites. *Genome Biol*. 2015; 16:191. doi: [10.1186/s13059-015-0759-1](https://doi.org/10.1186/s13059-015-0759-1) PMID: [26374288](https://pubmed.ncbi.nlm.nih.gov/26374288/)
44. Chaston JM, Dobson AJ, Newell PD, Douglas AE. Host genetic control of the microbiota mediates the *Drosophila* nutritional phenotypes. *Appl Environ Microbiol*. 2016; 82:671–9.
45. Obbard DJ, Welch JJ, Kim K-W, Jiggins FM. Quantifying adaptive evolution in the *Drosophila* immune system. *PLoS Genet*. 2009; 5:e1000698. doi: [10.1371/journal.pgen.1000698](https://doi.org/10.1371/journal.pgen.1000698) PMID: [19851448](https://pubmed.ncbi.nlm.nih.gov/19851448/)
46. Nielsen R, Bustamante C, Clark AG, Glanowski S, Sackton TB, Hubisz MJ, et al. A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol*. 2005; 3:e170. doi: [10.1371/journal.pbio.0030170](https://doi.org/10.1371/journal.pbio.0030170) PMID: [15869325](https://pubmed.ncbi.nlm.nih.gov/15869325/)
47. Hooper LV, Gordon JI. Glycans as legislators of host-microbial interactions: spanning the spectrum from symbiosis to pathogenicity. *Glycobiol*. 2001; 11:1–10.
48. McLoughlin K, Schluter J, Rakoff-Nahoum S, Smith AL, Foster KR. Host selection of microbiota via differential adhesion. *Cell Host & Microbe*. 2016; 19(4):550–9 doi: <http://dx.doi.org/10.1016/j.chom.2016.02.021>.
49. Kashyap PC, Macobal A, Ursell LK, Smits SA, Sonnenburg ED, Costello EK, et al. Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota. *Proc Natl Acad Sci*. 2013; 110:17059–64. doi: [10.1073/pnas.1306070110](https://doi.org/10.1073/pnas.1306070110) PMID: [24062455](https://pubmed.ncbi.nlm.nih.gov/24062455/)
50. Liang X, Bushman FD, FitzGerald GA. Rhythmicity of the intestinal microbiota is regulated by gender and the host circadian clock. *Proc Natl Acad Sci*. 2015; 112:10479–84. doi: [10.1073/pnas.1501305112](https://doi.org/10.1073/pnas.1501305112) PMID: [26240359](https://pubmed.ncbi.nlm.nih.gov/26240359/)
51. Liu S, Pires da Cunha A, Rezende RM, Cialic R, Wei Z, Bry L, et al. The host shapes the gut microbiota via fecal microRNA. *Cell Host & Microbe*. 2016; 19:32–43.
52. Malo MS, Alam SN, Mostafa G, Zeller SJ, Johnson PV, Mohammad N, et al. Intestinal alkaline phosphatase preserves the normal homeostasis of gut microbiota. *Gut*. 2010; 59:1476–84. doi: [10.1136/gut.2010.211706](https://doi.org/10.1136/gut.2010.211706) PMID: [20947883](https://pubmed.ncbi.nlm.nih.gov/20947883/)
53. Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA. The application of ecological theory toward an understanding of the human microbiome. *Science*. 2012; 336(1255–1262). doi: [10.1126/science.1224203](https://doi.org/10.1126/science.1224203) PMID: [22674335](https://pubmed.ncbi.nlm.nih.gov/22674335/)
54. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Loso T, Douglas AE, et al. Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci*. 2013; 110:3229–36. doi: [10.1073/pnas.1218525110](https://doi.org/10.1073/pnas.1218525110) PMID: [23391737](https://pubmed.ncbi.nlm.nih.gov/23391737/)
55. Seedorf H, Griffin NW, Ridaura VK, Reyes A, Cheng J, Rey FE, et al. Bacteria from diverse habitats colonize and compete in the mouse gut. *Cell*. 2014; 159:253–66. doi: [10.1016/j.cell.2014.09.008](https://doi.org/10.1016/j.cell.2014.09.008) PMID: [25284151](https://pubmed.ncbi.nlm.nih.gov/25284151/)
56. van Opstal EJ, Bordenstein SR. Rethinking heritability of the microbiome. *Science*. 2015; 349:1172–3. doi: [10.1126/science.aab3958](https://doi.org/10.1126/science.aab3958) PMID: [26359393](https://pubmed.ncbi.nlm.nih.gov/26359393/)
57. Funkhouser LJ, Bordenstein SR. Mom knows best: the universality of maternal microbial transmission. *PLoS Biol*. 2013; 11:e1001631. doi: [10.1371/journal.pbio.1001631](https://doi.org/10.1371/journal.pbio.1001631) PMID: [23976878](https://pubmed.ncbi.nlm.nih.gov/23976878/)
58. Sharp KH, Eam B, Faulkner DJ, Haygood MG. Vertical transmission of diverse microbes in the tropical sponge *Corticium* sp. *Appl Environ Microbiol*. 2007; 73:622–9. doi: [10.1128/AEM.01493-06](https://doi.org/10.1128/AEM.01493-06) PMID: [17122394](https://pubmed.ncbi.nlm.nih.gov/17122394/)
59. Pantoja-Feliciano IG, Clemente JC, Costello EK, Perez ME, Blaser MJ, Knight R, et al. Biphasic assembly of the murine intestinal microbiota during development. *ISME Journal*. 2013; 7:1112–5. doi: [10.1038/ismej.2013.15](https://doi.org/10.1038/ismej.2013.15) PMID: [23535917](https://pubmed.ncbi.nlm.nih.gov/23535917/)
60. Yatsunenkov T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. *Nature*. 2012; 486:222–7. doi: [10.1038/nature11053](https://doi.org/10.1038/nature11053) PMID: [22699611](https://pubmed.ncbi.nlm.nih.gov/22699611/)
61. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, et al. Human genetics shape the gut microbiome. *Cell*. 2014; 159:789–99. doi: [10.1016/j.cell.2014.09.053](https://doi.org/10.1016/j.cell.2014.09.053) PMID: [25417156](https://pubmed.ncbi.nlm.nih.gov/25417156/)

62. Agler MT, Ruhe J, Kroll S, Morhenn C, Kim ST, Weigel D, et al. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS Biol.* 2016; 14:e1002352. doi: [10.1371/journal.pbio.1002352](https://doi.org/10.1371/journal.pbio.1002352) PMID: [26788878](https://pubmed.ncbi.nlm.nih.gov/26788878/)
63. Fisher CK, Mehta P. Identifying keystone species in the human gut microbiome from metagenomic timeseries using sparse linear regression. *PLoS ONE.* 2014; 9:e102451. doi: [10.1371/journal.pone.0102451](https://doi.org/10.1371/journal.pone.0102451) PMID: [25054627](https://pubmed.ncbi.nlm.nih.gov/25054627/)
64. Murfin KE, Lee M- M, Klassen JL, McDonald BR, Larget B, Forst S, et al. *Xenorhabdus bovienii* strain diversity impacts coevolution and symbiotic maintenance with *Steinernema* spp. nematode hosts. *mBio.* 2015; 6:e00076–15. doi: [10.1128/mBio.00076-15](https://doi.org/10.1128/mBio.00076-15) PMID: [26045536](https://pubmed.ncbi.nlm.nih.gov/26045536/)
65. Wang J, Kalyan S, Steck N, Turner LM, Harr B, Künzel S, et al. Analysis of intestinal microbiota in hybrid house mice reveals evolutionary divergence in a vertebrate hologenome. *Nat Comm.* 2015; 6:6440.
66. Boot R, Koopman JP, Kruijt BC, Lammers RM, Kennis HM, Lankhorst A, et al. The 'normalization' of germ-free rabbits with host-specific caecal microflora. *Lab Anim.* 1985; 19:344–52. PMID: [4068663](https://pubmed.ncbi.nlm.nih.gov/4068663/)
67. Wostmann BS. Germfree and gnotobiotic animal models: background and applications. Boca Raton, FL: CRC Press; 1996.
68. Chung H, Pamp SJ, Hill JA, Surana NK, Edelman SM, Troy EB, et al. Gut immune maturation depends on colonization with a host-specific microbiota. *Cell.* 2012; 149:1578–93. doi: [10.1016/j.cell.2012.04.037](https://doi.org/10.1016/j.cell.2012.04.037) PMID: [22726443](https://pubmed.ncbi.nlm.nih.gov/22726443/)
69. Alegado RA, Brown LW, Cao S, Dermenjian RK, Zuzow R, Fairclough SR, et al. A bacterial sulfonolipid triggers multicellular development in the closest living relatives of animals. *eLife.* 2012; 1:e00013. doi: [10.7554/eLife.00013](https://doi.org/10.7554/eLife.00013) PMID: [23066504](https://pubmed.ncbi.nlm.nih.gov/23066504/)
70. Roth O, Sadd BM, Schmid-Hempel P, Kurtz J. Strain-specific priming of resistance in the red flour beetle, *Tribolium castaneum*. *Proc R Soc B.* 2009; 276:145–51. doi: [10.1098/rspb.2008.1157](https://doi.org/10.1098/rspb.2008.1157) PMID: [18796392](https://pubmed.ncbi.nlm.nih.gov/18796392/)
71. Bordenstein SR, Theis KR. Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS Biol.* 2015; 13:e1002226. doi: [10.1371/journal.pbio.1002226](https://doi.org/10.1371/journal.pbio.1002226) PMID: [26284777](https://pubmed.ncbi.nlm.nih.gov/26284777/)
72. Shapira M. Gut microbiotas and host evolution: scaling up symbiosis. *Trends Ecol Evol.* 2016. E-pub ahead of print. doi: <http://dx.doi.org/10.1016/j.tree.2016.03.006>
73. Theis KR, Dheilly NM, Klassen JL, Brucker RM, Baines JF, Bosch TCG, et al. Getting the hologenome concept right: An eco-evolutionary framework for hosts and their microbiomes. *mSystems.* 2016; 1:e00028–16.
74. Moeller AH, Caro-Quintero A, Mjunga D, Georgiev AV, Lonsdorf EV, Muller MN, et al. Cospeciation of gut microbiota with hominids. *Science.* 2016; 353:380–2. doi: [10.1126/science.aaf3951](https://doi.org/10.1126/science.aaf3951) PMID: [27463672](https://pubmed.ncbi.nlm.nih.gov/27463672/)
75. Nishiguchi MK. Host-symbiont recognition in the environmentally transmitted sepiolid squid-Vibrio mutualism. *Microb Ecol.* 2002; 44:10–8. doi: [10.1007/s00248-002-0002-y](https://doi.org/10.1007/s00248-002-0002-y) PMID: [12019463](https://pubmed.ncbi.nlm.nih.gov/12019463/)
76. Lo N, Casiraghi M, Salati E, Bazzocchi C, Bandi C. How many *Wolbachia* supergroups exist? *Mol Biol Evol.* 2002; 19:341–6. PMID: [11861893](https://pubmed.ncbi.nlm.nih.gov/11861893/)
77. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Method.* 2010; 7:335–6.
78. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics.* 2011; 27:2194–200. doi: [10.1093/bioinformatics/btr381](https://doi.org/10.1093/bioinformatics/btr381) PMID: [21700674](https://pubmed.ncbi.nlm.nih.gov/21700674/)
79. McDonald D, Price MN, Goodrich JK, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved GreenGenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal.* 2012; 6:610–8. doi: [10.1038/ismej.2011.139](https://doi.org/10.1038/ismej.2011.139) PMID: [22134646](https://pubmed.ncbi.nlm.nih.gov/22134646/)
80. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics.* 2010; 26:2460–1. doi: [10.1093/bioinformatics/btq461](https://doi.org/10.1093/bioinformatics/btq461) PMID: [20709691](https://pubmed.ncbi.nlm.nih.gov/20709691/)
81. Russell DJ, Otu HH, Sayood K. Grammar-based distance in progressive multiple sequence alignment. *BMC Bioinformatics.* 2008; 9:306. doi: [10.1186/1471-2105-9-306](https://doi.org/10.1186/1471-2105-9-306) PMID: [18616828](https://pubmed.ncbi.nlm.nih.gov/18616828/)
82. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE.* 2010; 5:e9490. doi: [10.1371/journal.pone.0009490](https://doi.org/10.1371/journal.pone.0009490) PMID: [20224823](https://pubmed.ncbi.nlm.nih.gov/20224823/)
83. McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D, et al. The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *Giga-Science.* 2010; 1:7.
84. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc.* 2007; 2:2366–82. doi: [10.1038/nprot.2007.324](https://doi.org/10.1038/nprot.2007.324) PMID: [17947979](https://pubmed.ncbi.nlm.nih.gov/17947979/)

85. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*. 2013; 8:e61217. doi: [10.1371/journal.pone.0061217](https://doi.org/10.1371/journal.pone.0061217) PMID: [23630581](https://pubmed.ncbi.nlm.nih.gov/23630581/)
86. Patwardhan A, Ray S, Roy A. Molecular markers in phylogenetic studies—a review. *J Phylogen Evol Biol*. 2014; 2:131.
87. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acid Res*. 2004; 32:1792–7.
88. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Method*. 2012; 9:772.
89. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014; 30:1312–3. doi: [10.1093/bioinformatics/btu033](https://doi.org/10.1093/bioinformatics/btu033) PMID: [24451623](https://pubmed.ncbi.nlm.nih.gov/24451623/)
90. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 2010; 59:307–21. doi: [10.1093/sysbio/syq010](https://doi.org/10.1093/sysbio/syq010) PMID: [20525638](https://pubmed.ncbi.nlm.nih.gov/20525638/)
91. Bogdanowicz D, Giaro K, Wróbel B. TreeCmp: comparison of trees in polynomial time. *Evol Bioinform Online*. 2012; 8:475.
92. Millman KJ, Aivazis M. Python for scientists and engineers. *Comp Sci Eng*. 2011; 13:9–12.
93. Oliphant TE. Python for scientific computing. *Comp Sci Eng*. 2007; 9:10–20.
94. Meagher S, Penn D, Potts W. Male-male competition magnifies inbreeding depression in wild house mice. *Proc Natl Acad Sci*. 2000; 97:3324–9. doi: [10.1073/pnas.060284797](https://doi.org/10.1073/pnas.060284797) PMID: [10716731](https://pubmed.ncbi.nlm.nih.gov/10716731/)
95. Gaukler SM, Ruff JS, Galland T, Underwood TK, Kandaris KA, Liu NM, et al. Quantification of cerivastatin toxicity supports organismal performance assays as an effective tool during pharmaceutical safety assessment. *Evol Appl*. 2016; 9(5):685–96. doi: [10.1111/eva.12365](https://doi.org/10.1111/eva.12365) PMID: [27247619](https://pubmed.ncbi.nlm.nih.gov/27247619/)
96. Zhang C, Derrien M, Levenez F, Brazeilles R, Ballal S, Kim J, et al. Ecological robustness of the gut microbiota in response to the ingestion of transient food-borne microbes. *ISME Journal*. 2016; 10:2235–45. doi: [10.1038/ismej.2016.13](https://doi.org/10.1038/ismej.2016.13) PMID: [26953599](https://pubmed.ncbi.nlm.nih.gov/26953599/)