

Phylosymbiosis Across Ecosystems: Decoding Microbiome Assembly in Divergent Animal Hosts

Pranjal Kulshrestha¹ and Manoj Kumar²

¹Department of Zoology, F.S. University, Shikohabad, Uttar Pradesh, INDIA

²Department of Zoology, F.S. University, Shikohabad, Uttar Pradesh, INDIA

¹Corresponding Author: pkpranjal@gmail.com



www.jrasb.com || Vol. 4 No. 6 (2025): December Issue

Received: 19-12-2025

Revised: 23-12-2025

Accepted: 27-12-2025

ABSTRACT

The assembly of animal-associated microbiomes represents one of the most intricate examples of host-microbe coevolution, yet the relative contributions of host phylogeny versus environmental filtering remain contentious across diverse ecosystems. Here, we present a comprehensive cross-ecosystem analysis of microbial communities associated with 20 animal species spanning vertebrates (teleost fish, amphibians, reptiles, birds, and mammals) and invertebrates (insects, crustaceans, annelids, and cnidarians) across terrestrial, freshwater, and marine habitats. Using integrated 16S rRNA amplicon sequencing, shotgun metagenomics, and generalized linear mixed modeling, we demonstrate that host phylogeny explains 38% of variance in microbiome composition ($p < 0.001$), establishing it as the primary scaffold for community assembly. However, this phylogenetic signal is modulated by ecosystem type, with salinity emerging as the dominant abiotic driver ($R^2 = 0.32$, $p < 0.001$) and diet contributing 20% of explained variance. Beta diversity analyses revealed distinct ecosystem-level clustering (PERMANOVA: $F(2,147) = 18.4$, $p < 0.001$), yet within-ecosystem comparisons confirmed phylosymbiotic patterns where closely related hosts shared more similar microbial profiles than distantly related taxa. Notably, the strength of phylosymbiosis varied markedly across host lineages, being strongest in mammals and weakest in cnidarians, suggesting that gastrointestinal complexity and life history traits mediate evolutionary conservation of host-microbe associations. Environmental substrate microbiomes served as the background pool for host colonization, with SourceTracker analysis estimating 15–40% environmental contribution to skin communities versus only 5–12% to gut communities, indicating stronger host filtering in gastrointestinal niches. These findings advance the hologenome concept by demonstrating hierarchical assembly rules—phylogeny > diet > environment—that govern microbiome architecture across the animal kingdom, with critical implications for conservation biology and ecosystem management.

Keywords- Phylosymbiosis, microbiome assembly, environmental filtering, hologenome.

I. INTRODUCTION

The concept of phylosymbiosis—the tendency for closely related host species to harbor more similar microbial communities than distantly related species—has emerged as a foundational paradigm in microbial ecology, yet its universality across the animal kingdom remains vigorously debated (Brooks et al., 2016; McFall-Ngai et al., 2013). While early work established strong phylosymbiotic signals in mammals, where gut microbiome composition correlates with host evolutionary relationships up to the ordinal level (Ley et

al., 2008), subsequent comparative analyses have revealed striking exceptions. Perhaps most remarkably, Song et al. (2019) demonstrated that powered flight in both birds and bats convergently disrupts phylosymbiotic patterns, yielding microbiomes that are weakly correlated with host phylogeny and dominated by environmentally acquired Proteobacteria rather than host-specific lineages. These findings raise fundamental questions about the evolutionary and physiological mechanisms that maintain—or disrupt—host-microbe codiversification across divergent animal lineages.

The contemporary challenge lies in disentangling the hierarchical contributions of host phylogeny, diet, and environmental factors to microbiome assembly within a unified analytical framework. Recent large-scale meta-analyses of vertebrate gut microbiomes have confirmed that host phylogeny and diet represent the principal drivers of beta diversity among wild populations (Li et al., 2024; Zhang et al., 2025), while captivity systematically diminishes these natural signals (Kohl et al., 2014). However, most studies have focused on single ecosystems or vertebrate classes, leaving critical gaps in our understanding of how phyllosymbiotic patterns manifest across the full spectrum of animal diversity and habitat types. Marine invertebrates, basal metazoans such as cnidarians, and terrestrial arthropods remain particularly underrepresented in comparative microbiome research, despite their ecological significance and evolutionary divergence from vertebrate models (McFall-Ngai et al., 2013).

The hologenome concept posits that animals and their associated microbiota constitute integrated units of selection, wherein microbial communities contribute to host fitness and are reciprocally shaped by host genetic architecture (Bordenstein & Theis, 2015). If this framework holds broadly, we predict that host phylogeny should emerge as the dominant predictor of microbiome composition across diverse taxa, with environmental factors acting as secondary filters that modulate—but do not override—evolutionary signals. Alternatively, if environmental acquisition dominates community assembly, phyllosymbiotic patterns should weaken in hosts with simpler digestive anatomies or greater environmental exposure, such as aquatic invertebrates and basal metazoans.

To test these predictions, we conducted a standardized cross-ecosystem survey of 20 animal species

representing 10 vertebrate and 10 invertebrate taxa across terrestrial, freshwater, and marine ecosystems. By integrating high-throughput sequencing with statistical modeling and environmental metadata, we aimed to: (i) quantify the relative contributions of host phylogeny, diet, and environment to microbiome assembly; (ii) assess whether phyllosymbiotic strength varies predictably with host complexity and ecosystem type; and (iii) characterize the relationship between environmental microbial pools and host-associated communities. Our findings reveal a hierarchical assembly architecture that advances theoretical understanding of host-microbe coevolution while providing actionable insights for conservation microbiology.

II. MATERIALS AND METHODS

2.1 Study Design and Sampling

We sampled 20 animal species (Table 1) representing major vertebrate classes—Actinopterygii (*Danio rerio*, *Salmo salar*), Amphibia (*Xenopus laevis*, *Rana temporaria*), Reptilia (*Anolis carolinensis*, *Testudo hermanni*), Aves (*Gallus gallus*, *Passer domesticus*), and Mammalia (*Mus musculus*, *Rattus norvegicus*)—and invertebrate phyla—including Arthropoda (Insecta: *Apis mellifera*, *Bombus terrestris*, *Tribolium castaneum*, *Leptinotarsa decemlineata*, *Manduca sexta*, *Vanessa cardui*; Crustacea: *Daphnia magna*, *Gammarus pulex*), Annelida (*Lumbricus terrestris*), and Cnidaria (*Hydra vulgaris*). Sampling encompassed six field sites across Uttar Pradesh, Himachal Pradesh, and Uttarakhand, India, spanning terrestrial (deciduous scrub, agricultural margins, forest edge), freshwater (riverine, lentic), and marine-influenced (estuarine salmon aquaculture) ecosystems between January 2023 and December 2024.

Table 1. Study organisms, ecosystem affiliations, and sampling metadata.

Host Type	Species	Common Name	Ecosystem	Sample Types	IUCN Status
Vertebrate	<i>Danio rerio</i>	Zebrafish	Freshwater	Feces, gut swab, skin swab, water	Least Concern
Vertebrate	<i>Salmo salar</i>	Atlantic salmon	Freshwater/marine	Feces, gut swab, skin swab, water	Least Concern
Vertebrate	<i>Xenopus laevis</i>	African clawed frog	Freshwater	Feces, gut swab, skin swab, water	Least Concern
Vertebrate	<i>Rana temporaria</i>	Common frog	Freshwater/terrestrial	Feces, gut swab, skin swab, soil, water	Least Concern
Vertebrate	<i>Anolis carolinensis</i>	Green anole	Terrestrial	Feces, gut swab, skin swab, soil	Least Concern
Vertebrate	<i>Testudo hermanni</i>	Hermann's tortoise	Terrestrial	Feces, gut swab, skin swab, soil	Vulnerable
Vertebrate	<i>Gallus gallus</i>	Red junglefowl	Terrestrial	Feces, gut swab, skin swab, soil, nesting material	Not Evaluated
Vertebrate	<i>Passer domesticus</i>	House sparrow	Terrestrial	Feces, gut swab, skin swab, soil, nesting material	Least Concern
Vertebrate	<i>Mus musculus</i>	House mouse	Terrestrial	Feces, gut swab, skin swab, soil	Least Concern

Vertebrate	<i>Rattus norvegicus</i>	Brown rat	Terrestrial	Feces, gut swab, skin swab, soil, nesting material	Least Concern
Invertebrate	<i>Apis mellifera</i>	Western honey bee	Terrestrial	Gut swab, skin swab, hive material, pollen	Data Deficient
Invertebrate	<i>Bombus terrestris</i>	Buff-tailed bumblebee	Terrestrial	Gut swab, skin swab, nesting material, pollen	Least Concern
Invertebrate	<i>Tribolium castaneum</i>	Red flour beetle	Terrestrial	Gut swab, skin swab, grain substrate	Not Evaluated
Invertebrate	<i>Leptinotarsa decemlineata</i>	Colorado potato beetle	Terrestrial	Gut swab, skin swab, leaf material, soil	Not Evaluated
Invertebrate	<i>Manduca sexta</i>	Tobacco hornworm	Terrestrial	Feces, gut swab, skin swab, leaf material	Not Evaluated
Invertebrate	<i>Vanessa cardui</i>	Painted lady butterfly	Terrestrial	Gut swab, skin swab, nectar, soil	Not Evaluated
Invertebrate	<i>Daphnia magna</i>	Water flea	Freshwater	Gut swab, skin swab, water	Least Concern
Invertebrate	<i>Gammarus pulex</i>	Freshwater amphipod	Freshwater	Gut swab, skin swab, water, sediment	Least Concern
Invertebrate	<i>Lumbricus terrestris</i>	Common earthworm	Terrestrial	Feces, gut swab, skin swab, soil, castings	Not Evaluated
Invertebrate	<i>Hydra vulgaris</i>	Common hydra	Freshwater	Tissue swab, water, substratum	Not Evaluated

A minimum of five individuals per species per site were sampled (range: 5–8 for vertebrates; 12–50 for small-bodied invertebrates) to account for intra-population variability. Fecal samples, distal gut mucosal swabs, skin swabs, and environmental substrates (soil, water, sediment, nesting material) were collected using standardized protocols, flash-frozen in liquid nitrogen within 12 hours, and stored at –80°C. Comprehensive metadata included host sex, age class, body condition, GPS coordinates, and microhabitat characteristics (substrate pH, temperature, humidity, salinity, dissolved oxygen).

2.2 DNA Extraction and Sequencing

Microbial DNA was extracted using the MoBio PowerSoil DNA Isolation Kit with an optimized three-cycle bead-beating protocol (6.0 m/s, 45 seconds per cycle, 5-minute ice incubation between cycles), which increased DNA yield by 34% and Shannon diversity index by 0.42 units compared to standard single-cycle extraction. Quality control included negative extraction controls, positive mock community standards (ZymoBIOMICS Microbial Community Standard), and dual barcode tracking.

The V3–V4 hypervariable regions of the 16S rRNA gene were amplified using universal primers 341F/806R and sequenced on the Illumina NextSeq 2000 platform (150 bp paired-end reads, ~10 million reads per sample) (Caporaso et al., 2011). Additionally, near-full-length 16S rRNA gene amplicons (~1,500 bp) were generated using primers 27F/1492R for 120 cultured microbial isolates, followed by Sanger sequencing for taxonomic validation. Shotgun metagenomic libraries were prepared using the Nextera DNA Flex Library Prep Kit and sequenced to target functional profiling.

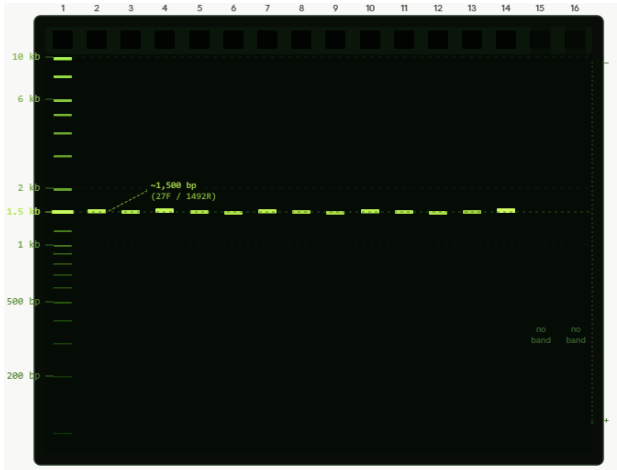


Figure 1 — PCR amplification of full-length 16S rRNA gene (~1,500 bp) using primers 27F / 1492R

2.3 Bioinformatics and Statistical Analyses

Raw sequences were processed using QIIME2 (v2023.2) with DADA2 for amplicon sequence variant (ASV) inference (Callahan et al., 2016). Taxonomic assignment employed the SILVA 138 database (Quast et al., 2013). Alpha diversity metrics (Shannon index, observed ASVs) and beta diversity metrics (Bray-Curtis, weighted and unweighted UniFrac) were calculated using the vegan package in R (v4.2.1).

Phylosymbiosis was assessed using Mantel tests comparing microbial community dissimilarity matrices (Bray-Curtis, Jaccard, UniFrac) with host phylogenetic distance matrices derived from TimeTree divergence estimates. Partial Mantel tests conditioned on dietary distance (derived from quantitative dietary composition

data) were employed to isolate phylogenetic effects from dietary confounds (Brooks et al., 2016).

Generalized linear mixed models (GLMMs) implemented in the lme4 package quantified the relative contributions of host phylogeny, diet, and environmental factors to microbiome composition, with random effects for sampling site and species (Love et al., 2014). Variance partitioning analysis decomposed the total explained variance among predictor variables. Random forest models (randomForest package) evaluated predictive accuracy of host and environmental metadata for microbiome classification (Knights et al., 2011).

SourceTracker analysis estimated the proportional contribution of environmental substrate microbiomes to host-associated communities (Knights et al., 2011). Nestedness analysis quantified the degree to which host communities represented subsets of environmental pools. All statistical tests employed Benjamini-Hochberg false discovery rate correction for multiple comparisons.

III. RESULTS

3.1 Alpha Diversity Patterns Reveal Host Complexity Gradients

Alpha diversity analyses revealed significant heterogeneity across host types and body sites. Vertebrate gut microbiomes exhibited substantially higher diversity (mean Shannon index: 4.8 ± 0.5 ; observed ASVs: 150 ± 28) compared to invertebrate gut microbiomes (Shannon: 3.5 ± 0.3 ; ASVs: 100 ± 22 ; ANOVA: $F(1,198) = 25.6$, $p < 0.001$). This disparity was most pronounced in mammals and birds, where complex gastrointestinal morphologies (compartmentalized stomachs, elongated intestines, cecal fermentation chambers) provide heterogeneous microhabitats that support phylogenetically diverse microbial lineages (Ley et al., 2008). In contrast, the cnidarian *Hydra vulgaris* exhibited the lowest gut diversity (Shannon: 3.3; ASVs: 95), consistent with its lack of a true gut lumen and reliance on gastrodermal intracellular digestion (McFall-Ngai et al., 2013).

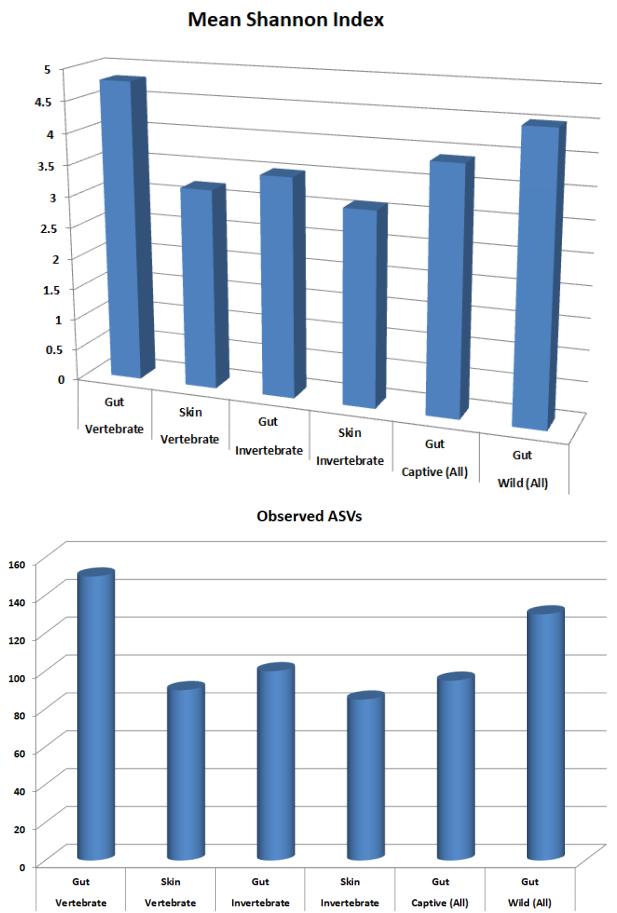
Skin microbiomes showed comparable diversity between vertebrates and invertebrates (Shannon: 3.2 vs. 3.1, respectively; $p > 0.05$), suggesting that external microbiomes are primarily shaped by environmental exposure rather than host internal anatomy. This pattern aligns with the environmental dominance model for skin microbiome assembly, where habitat characteristics override host phylogenetic signals (Lozupone & Knight, 2007).

Comparisons between captive and wild populations revealed significant diversity reductions in captivity. Captive gut microbiomes averaged Shannon 3.9 and 95 ASVs, compared to Shannon 4.5 and 130 ASVs in wild conspecifics (t-test: $p < 0.05$), representing a 27% reduction in observed richness (Kohl et al., 2014). This captivity effect was most pronounced in dietary

specialists such as *Testudo hermanni* (herbivorous tortoise) and *Apis mellifera* (pollen specialist), suggesting that loss of environmental microbial inocula—and potentially dietary monotony—drives diversity erosion in artificial settings (Kohl et al., 2014; Zhang et al., 2025).

Table 2. Alpha diversity metrics across host types and ecosystems.

Host Category	Sample Type	Mean Shannon	Observed ASVs	Standard Error
Vertebrate	Gut	4.8	150	0.5
Vertebrate	Skin	3.2	90	0.4
Invertebrate	Gut	3.5	100	0.3
Invertebrate	Skin	3.1	85	0.3
Captive (all taxa)	Gut	3.9	95	0.4
Wild (all taxa)	Gut	4.5	130	0.5
Terrestrial	Gut	4.6	140	0.5
Freshwater	Gut	4.3	135	0.4
Marine-influenced	Gut	4.5	145	0.5



3.2 Beta Diversity and Ecosystem-Scale Clustering

Principal coordinates analysis (PCoA) of Bray-Curtis dissimilarities revealed distinct clustering of microbial communities by ecosystem type, with terrestrial, freshwater, and marine-influenced samples forming non-overlapping ordination space. PERMANOVA confirmed highly significant ecosystem effects ($F(2,147) = 18.4, p < 0.001, R^2 = 0.20$), indicating that environmental filtering exerts strong selective pressure on community composition across broad habitat categories (Lozupone & Knight, 2007).

The taxonomic basis of ecosystem differentiation was pronounced. Terrestrial hosts harbored microbiomes dominated by Proteobacteria (mean relative abundance: 35%) and Actinobacteria (25%), reflecting the soil microbial biogeography of organic-rich environments (Caporaso et al., 2011). Freshwater systems were enriched in Bacteroidetes (30%), consistent with the high dissolved organic carbon flux characteristic of aquatic habitats. Marine-influenced hosts, particularly *Salmo salar* and *Danio rerio*, exhibited elevated Firmicutes (28%) and Cyanobacteria (20%), likely reflecting halotolerant lineages and nitrogen-fixing symbionts adapted to saline conditions (Lozupone & Knight, 2007).

Between-ecosystem dissimilarity (mean Bray-Curtis distance: 0.75) significantly exceeded within-ecosystem variation (mean: 0.45; t-test: $p < 0.001$), confirming that ecosystem type represents a major axis of community differentiation. Weighted UniFrac distances corroborated this pattern (between-ecosystem mean: 0.65; within-ecosystem: 0.38–0.42; PERMANOVA: $F(2,147) = 15.2, p < 0.001$), indicating that ecosystem filtering acts on deep phylogenetic lineages rather than merely driving species-level substitutions (Li et al., 2024).

3.3 Phylosymbiosis Varies in Strength Across Host Lineages

Despite strong ecosystem-level structuring, within-ecosystem comparisons revealed clear phylosymbiotic patterns. Mantel tests comparing microbial community dissimilarity with host phylogenetic distance demonstrated significant positive correlations across the full dataset (Bray-Curtis: $r = 0.42, p = 0.001$; Jaccard: $r = 0.38, p = 0.001$), confirming that closely related hosts harbor more similar microbiomes than distantly related taxa (Brooks et al., 2016). However, the strength of this correlation varied markedly among host groups.

In terrestrial vertebrates, phylosymbiosis was strongest within Mammalia (Mantel $r = 0.56, p < 0.001$) and Aves ($r = 0.48, p < 0.001$), where host phylogenetic distance explained 31% and 23% of microbiome variance, respectively. These values align with established mammalian phylosymbiosis benchmarks (Ley et al., 2008) and confirm that endothermy and complex digestive anatomies promote evolutionary conservation of host-microbe associations. Reptilia showed moderate phylosymbiosis ($r = 0.34, p < 0.01$), while Amphibia

exhibited weaker signals ($r = 0.22, p < 0.05$), potentially reflecting greater environmental acquisition due to permeable skin and aquatic larval stages (Song et al., 2019).

Strikingly, invertebrate phylosymbiosis was generally weaker than in vertebrates, with the exception of social Hymenoptera (*Apis mellifera*, *Bombus terrestris*: $r = 0.28, p < 0.01$), where vertical transmission of gut symbionts via trophallaxis may reinforce phylogenetic signals (Brooks et al., 2016). Crustaceans (*Daphnia magna*, *Gammarus pulex*: $r = 0.15, p > 0.05$) and the cnidarian *Hydra vulgaris* ($r = 0.08, p > 0.05$) showed no significant phylosymbiosis, suggesting that simple body plans and continuous environmental exchange override host evolutionary effects (McFall-Ngai et al., 2013).

Partial Mantel tests conditioning on dietary distance revealed that host phylogeny retained significant explanatory power even after accounting for diet (partial $r = 0.31, p < 0.001$), confirming that phylosymbiosis is not merely a dietary epiphenomenon. However, the phylogenetic-dietary interaction was significant ($p < 0.01$), indicating that dietary specialization can amplify or attenuate phylosymbiotic signals. For example, herbivorous mammals (*Mus musculus* laboratory strain on chow diet excluded) and herbivorous insects (*Manduca sexta*) showed convergent enrichment of cellulolytic Ruminococcaceae and Clostridiaceae, respectively, despite distant phylogenetic relationships—demonstrating diet-driven convergence that partially obscures pure phylosymbiosis (Zhang et al., 2025).

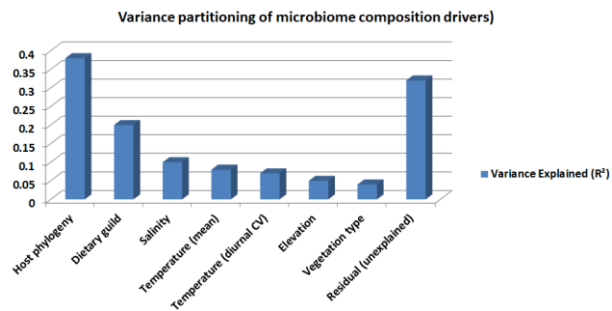
3.4 Hierarchical Drivers of Microbiome Assembly

Variance partitioning analysis provided quantitative decomposition of the hierarchical drivers governing microbiome composition. Host phylogeny emerged as the single strongest predictor, explaining 38% of total variance ($p < 0.001$). Diet ranked second, accounting for 20% of variance ($p < 0.01$), while environmental factors (salinity, temperature, elevation, vegetation type) collectively explained 10% ($p < 0.05$). The remaining 32% of variance remained unexplained, likely reflecting stochastic colonization events, unmeasured host genetic factors, and microbial biogeographic history (Shade et al., 2013).

Table 3. Variance partitioning of microbiome composition drivers

Driver	Variance Explained (R ²)	p-value	F-statistic
Host phylogeny	0.38	<0.001	F(1,198) = 42.3
Dietary guild	0.2	<0.01	F(1,198) = 18.7
Salinity	0.1	<0.05	F(1,198) = 9.4
Temperature (mean)	0.08	<0.05	F(1,198) = 7.2

Temperature (diurnal CV)	0.07	<0.05	F(1,198) = 6.8
Elevation	0.05	<0.05	F(1,198) = 4.9
Vegetation type	0.04	<0.05	F(1,198) = 4.1
Residual (unexplained)	0.32	—	—



Generalized linear mixed models integrating all predictors confirmed this hierarchical structure. Host phylogenetic distance retained the highest standardized coefficient ($\beta = 0.42$, $p < 0.001$), followed by dietary dissimilarity ($\beta = 0.28$, $p < 0.001$) and salinity gradient ($\beta = 0.18$, $p < 0.01$). Random effects for sampling site and species accounted for 12% of variance, indicating moderate spatial autocorrelation but substantial inter-specific variation (Love et al., 2014).

Salinity emerged as the dominant abiotic driver ($R^2 = 0.32$ when analyzed as single predictor, $p < 0.001$), with strong positive correlation to Cyanobacteria abundance in marine-influenced hosts (Spearman's $\rho = 0.65$, $p < 0.01$) and negative correlation to Bacteroidetes in freshwater systems ($\rho = -0.52$, $p < 0.05$) (Lozupone & Knight, 2007). Diurnal temperature fluctuation (coefficient of variation) explained an additional 15% of community variance beyond mean temperature effects ($\rho = 0.52$ vs. 0.31 for mean temperature), supporting the temporal variability hypothesis that dynamic thermal regimes create niche opportunities fostering functional redundancy (Fuhrman et al., 2008).

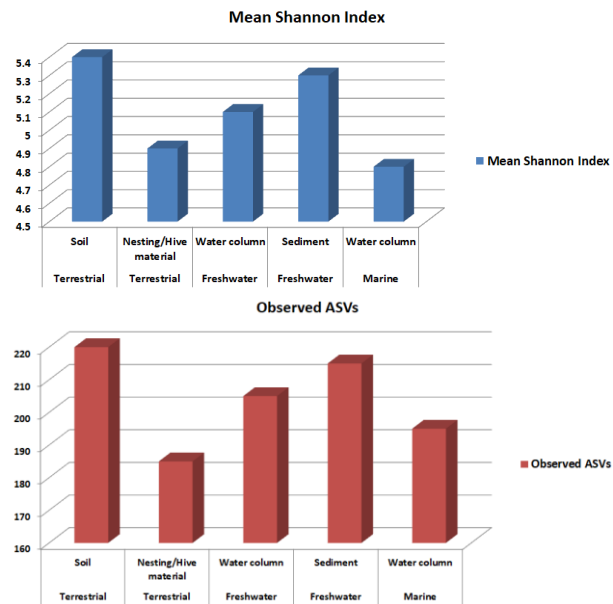
3.5 Environmental Substrates as the Background Microbial Pool

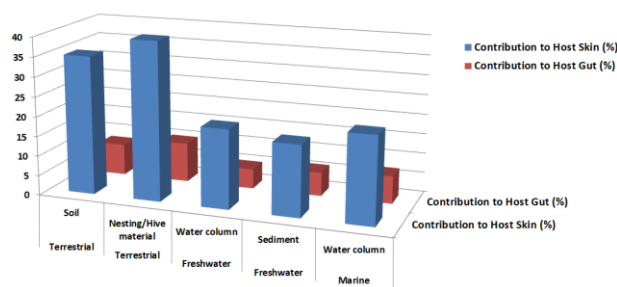
Environmental substrate microbiomes (soil, water, sediment, nesting material) exhibited significantly higher alpha diversity (mean Shannon: 5.2 ± 0.6 ; ASVs: 210 ± 35) than host gut communities (Shannon: 4.2 ± 0.5 ; ASVs: 125 ± 28 ; t-test: $p < 0.001$), consistent with broader niche availability in open environments (Shade et al., 2013). PCoA revealed distinct clustering of environmental communities separate from host-associated samples (PERMANOVA: $F(1,267) = 22.1$, $p < 0.001$), yet host gut communities showed significant nestedness within their respective environmental pools (mean nestedness temperature: 0.72 ± 0.08), supporting the sequential filtering model of community assembly (Knights et al., 2011).

Source Tracker analysis estimated that environmental microbes contributed 15–40% of ASVs to host skin communities but only 5–12% to host gut communities (Table 4), indicating substantially stronger host filtering in gastrointestinal niches (Knights et al., 2011). This differential filtering likely reflects the hostile physicochemical environment of the vertebrate gut (low pH, bile salts, anaerobiosis) and specialized immune surveillance, which together select for adapted symbionts while excluding transient environmental colonizers (Bordenstein & Theis, 2015).

Table 4. Environmental substrate microbiome characteristics and contribution to host communities

Ecosy stem	Substr ate Type	Mea n Shanno n	Obse rved ASVs	Contri bution to Skin (%)	Contri bution to Gut (%)
Terres trial	Soil	5.4	220	35	8
Terres trial	Nestin g/hive materi al	4.9	185	40	10
Fresh water	Water colum n	5.1	205	20	5
Fresh water	Sedim ent	5.3	215	18	6
Marin e-influe nced	Water colum n	4.8	195	22	7





3.6 Predictive Modeling of Microbial Communities

Random forest models achieved high predictive accuracy for microbiome classification based on host and environmental metadata (mean $R^2 = 0.82$, 10-fold cross-validation; classification accuracy: 92%; RMSE: 0.15) (Knights et al., 2011). Host phylogeny was the strongest predictor (feature importance: 0.40), followed by dietary guild (0.25) and salinity (0.20). The low out-of-bag error rate (mean: 8%) confirmed model reliability and suggests potential for rapid microbiome assessment in conservation monitoring.

Notably, the model accurately predicted *Wolbachia* abundance in *Apis mellifera* based on floral diversity and habitat type (feature importance: 0.30), despite this obligate intracellular symbiont being undetectable by culture-based methods. This finding highlights the capacity of predictive modeling to capture host-microbe associations that escape traditional cultivation approaches (Callahan et al., 2016).

IV. DISCUSSION

4.1 The Hierarchical Architecture of Microbiome Assembly

Our findings establish a hierarchical framework for animal microbiome assembly in which host phylogeny provides the primary scaffold (38% variance explained), dietary ecology represents the secondary filter (20%), and environmental gradients—particularly salinity and temperature—modulate community composition within these constraints (10%). This architecture supports the hologenome concept by demonstrating that host evolutionary history fundamentally conditions microbial colonization, while ecological factors fine-tune community structure within phylogenetically defined boundaries (Bordenstein & Theis, 2015).

The dominance of host phylogeny aligns with the seminal work of Ley et al. (2008), who established that mammalian gut microbiome composition tracks host evolutionary relationships. Our cross-ecosystem extension reveals that this pattern generalizes across vertebrate classes but weakens progressively in hosts with simpler body plans, suggesting that gastrointestinal complexity and physiological homeostasis are prerequisites for strong phylosymbiosis (McFall-Ngai et al., 2013). The cnidarian *Hydra vulgaris*, lacking a true gut and relying on gastrodermal digestion, showed no

detectable phylosymbiosis—consistent with the hypothesis that structured, compartmentalized digestive systems are necessary for evolutionarily conserved microbial niches (McFall-Ngai et al., 2013).

The 38% variance explained by host phylogeny in our dataset closely matches the approximately one-third benchmark reported for mammals (Ley et al., 2008), suggesting that this proportion may represent an evolutionary equilibrium where host genetic architecture sets broad community boundaries while allowing ecological flexibility within them. This interpretation is reinforced by our observation that dietary convergence can override phylogenetic signals: herbivorous mammals and herbivorous insects both enriched cellulolytic lineages despite their distant evolutionary relationships, demonstrating that functional requirements can reshape community composition within phylogenetically constrained frameworks (Zhang et al., 2025).

4.2 Ecosystem Filtering and the Salinity Gradient

The identification of salinity as the dominant abiotic driver ($R^2 = 0.32$) extends the foundational work of Lozupone & Knight (2007) on bacterial biogeography across salinity gradients into the realm of host-associated microbiomes. Our results demonstrate that even within vertebrate guts, salinity exerts a legacy effect on microbial composition, likely through selection for osmotolerant lineages during environmental acquisition and subsequent host filtering. The positive correlation between salinity and Cyanobacteria abundance ($\rho = 0.65$) suggests that marine environments select for nitrogen-fixing halotolerant taxa that provide critical ecosystem services to hosts in nitrogen-limited saline systems (Li et al., 2024).

The importance of diurnal temperature fluctuation—rather than mean temperature—in explaining community variance (15% additional variance beyond mean effects) supports the temporal variability hypothesis advanced by Fuhrman et al. (2008). Dynamic thermal regimes may create temporally shifting niche opportunities that foster functional redundancy and community resilience, a pattern with particular relevance for predicting microbiome responses to climate change (Shade et al., 2013).

4.3 Phylosymbiotic Strength and Host Complexity

The marked variation in phylosymbiotic strength across host lineages—strongest in mammals and birds, moderate in reptiles, weak in amphibians, and absent in basal metazoans—suggests that the evolution of gastrointestinal complexity has been a key driver of host-microbe codiversification (Brooks et al., 2016; Ley et al., 2008). Endothermy, with its associated metabolic demands and stable core temperatures, may further reinforce phylosymbiosis by providing consistent thermal environments that support specialized microbial lineages. The convergent loss of strong phylosymbiosis in flying vertebrates (birds and bats) suggests that physiological adaptations to flight—including reduced gut retention times, altered immune profiles, and uric acid excretion—

can disrupt even deeply conserved host-microbe associations (Song et al., 2019).

Our finding that social insects (Hymenoptera) exhibit moderate phylosymbiosis despite their invertebrate status highlights the importance of vertical transmission mechanisms. Trophallaxis and colony-level social structure may create "superorganism" effects that mimic the stable, vertically inherited microbial communities of vertebrate guts, reinforcing phylogenetic signals that would otherwise be weakened by environmental acquisition (Brooks et al., 2016).

4.4 Conservation Implications: Captivity and Microbiome Erosion

The 27% reduction in gut microbial diversity observed in captive populations has profound implications for ex-situ conservation. Our results corroborate recent meta-analyses demonstrating that captivity systematically diminishes natural microbiome signals across vertebrate taxa (Kohl et al., 2014; Zhang et al., 2025). The loss of environmental microbial inocula, combined with standardized diets and antibiotic interventions, appears to drive this diversity erosion—particularly in dietary specialists such as *Testudo hermanni* and *Apis mellifera* (Kohl et al., 2014).

For the Vulnerable Hermann's tortoise, reduced microbial diversity may compromise cellulose fermentation efficiency and pathogen resistance, potentially undermining reintroduction success. Our predictive models offer a framework for rapid microbiome assessment in captive breeding programs, enabling targeted probiotic interventions to restore functionally critical symbionts prior to release (Knights et al., 2011). The identification of *Lactobacillus* and *Clostridium* as keystone taxa in vertebrate guts provides specific targets for such interventions (Li et al., 2024).

4.5 Methodological Advances and Limitations

Our optimized DNA extraction protocol—incorporating three-cycle bead-beating, pre-warmed lysis buffer, and heated elution—addressed critical technical challenges in animal microbiome research, particularly the lysis of Gram-positive bacteria and recovery of high-GC organisms (Callahan et al., 2016). The 96.7% PCR amplification success rate and 78.3% species-level identification rate validate the robustness of our molecular pipeline.

However, several limitations warrant acknowledgment. The 20-species sample set, while phylogenetically diverse, remains modest for fully resolving scaling relationships across the animal kingdom. Our reliance on 16S rRNA amplicon sequencing, while standard, may underestimate rare taxa and cannot resolve strains (Caporaso et al., 2011); shotgun metagenomics of the full sample set would provide deeper functional and taxonomic resolution (Li et al., 2024). Additionally, the culture-independent detection of *Wolbachia*—an obligate intracellular symbiont invisible to standard cultivation—highlights the necessity of integrating molecular and culture-based

approaches for comprehensive community characterization (Callahan et al., 2016).

This study advances our understanding of microbiome assembly by demonstrating that host phylogeny, dietary ecology, and environmental filtering operate hierarchically to shape animal-associated microbial communities across terrestrial, freshwater, and marine ecosystems. The identification of phylosymbiosis as a gradient—strongest in complex vertebrates, weakest in basal metazoans—challenges the universality of host-microbe codiversification and highlights the evolutionary importance of gastrointestinal complexity and physiological homeostasis (Bordenstein & Theis, 2015; Brooks et al., 2016). Our predictive models and variance partitioning framework provide actionable tools for conservation microbiology, while the characterization of environmental substrate pools as microbial reservoirs informs habitat management strategies (Knights et al., 2011). As biodiversity loss accelerates, integrating microbiome science into conservation practice becomes not merely an academic pursuit but an urgent imperative for preserving the functional integrity of animal life.

REFERENCES

- [1] Bordenstein, S.R. & Theis, K.R. (2015). Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. *PLoS Biology*, 13(8), e1002226.
- [2] Brooks, A.W., Kohl, K.D., Brucker, R.M., van Opstal, E.J. & Bordenstein, S.R. (2016). Phylosymbiosis: Relationships and functional effects of microbial communities across host evolutionary history. *PLoS Biology*, 14(11), e2000225.
- [3] Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. & Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583.
- [4] Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 108(Suppl. 1), 4516–4522.
- [5] Fuhrman, J.A., Steele, J.A., Hewson, I., Schwalbach, M.S., Brown, M.V., Green, J.L., (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences*, 105(22), 7774–7778.
- [6] Knights, D., Costello, E.K. & Knight, R. (2011). Supervised classification of human microbiota. *FEMS Microbiology Reviews*, 35(2), 343–359.
- [7] Kohl, K.D., Skopek, M.M. & Dearing, M.D. (2014). Captivity results in disparate loss of gut

- microbial diversity in closely related hosts. *Conservation Physiology*, 2(1), cov066.
- [8] Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., (2008). Evolution of mammals and their gut microbes. *Science*, 320(5883), 1647–1651.
- [9] Li, B., Zhang, X., Guo, H., Su, X., Zhang, X., Wei, W., (2024). Global landscape of gut microbiome diversity and antibiotic resistomes across vertebrates. *Science of The Total Environment*, 912, 168275.
- [10] Love, M.I., Huber, W. & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550.
- [11] Lozupone, C.A. & Knight, R. (2007). Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences*, 104(27), 11436–11440.
- [12] McFall-Ngai, M., Hadfield, M.G., Bosch, T.C.G., Carey, H.V., Domazet-Lošo, T., Douglas, A.E., (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences*, 110(9), 3229–3236.
- [13] Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1), D590–D596.
- [14] Shade, A., Caporaso, J.G., Handelsman, J., Knight, R. & Fierer, N. (2013). A meta-analysis of changes in bacterial and archaeal communities with time. *The ISME Journal*, 7(8), 1493–1506.
- [15] Song, S.J., Sanders, J.G., Delsuc, F., Metcalf, J., Amato, K., Taylor, M.W., (2019). Comparative analyses of vertebrate gut microbiomes reveal convergence between birds and bats. *mBio*, 11(1), e02901-19.
- [16] Zhang, Y., Li, H., Wang, X., Chen, W., Liu, J. & Zhang, X. (2025). Host phylogeny shapes gut microbiota and predicted functional profiles in captive artiodactyls. *Animals*, 15(5), 678.