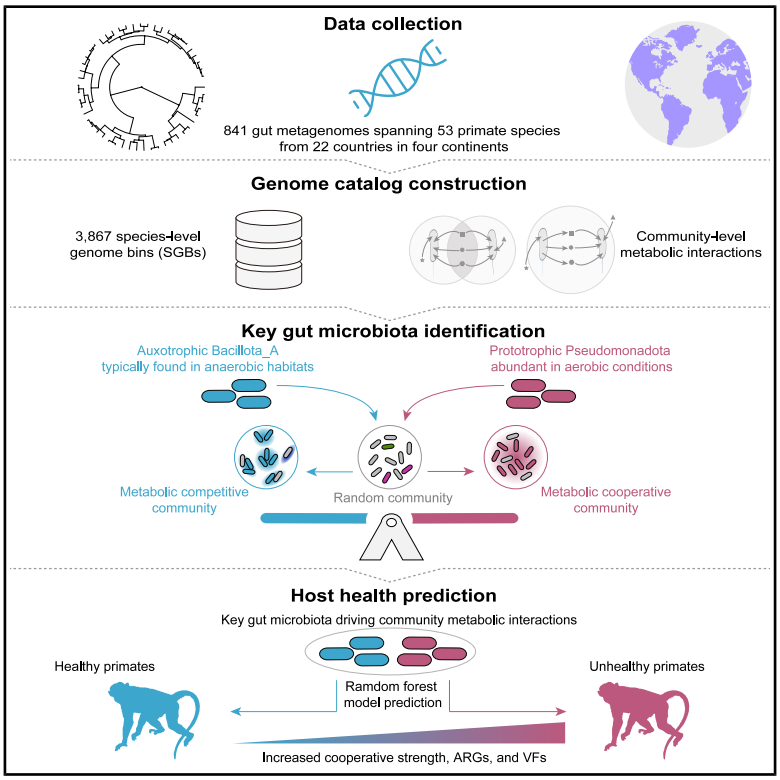
Cell Reports

Article

Health-associated key gut microbiota drives the variation in community metabolic interactions in non-human primates

Graphical abstract



Highlights

- A gut microbial genome catalog with 3,867 SGBs is constructed for 53 NHPs
- NHP phylogeny determines the MRO and MIP of gut microbial community
- Intrinsically competitive or cooperative key gut microbiota are identified
- Key competing and cooperative microbiota well predict the health status of NHPs

Authors

Shao-Ming Gao, Li-Ying Lan, Li Yang, Tao Chen, Peng-Fei Fan

Correspondence

fanpf@mail.sysu.edu.cn

In brief

Gao et al. reveal that the metabolic interactions of gut microbial community are driven by the key competitive and cooperative microbiota, which are potentially associated with non-human primate health. The study improves our understanding of the complex interactions in microbial community, providing potential avenues to therapeutic implications.

Gao et al., 2025, Cell Reports *44*, 116477
November 25, 2025
https://doi.org/10.1016/j.celrep.2025.116477

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Health-associated key gut microbiota drives the variation in community metabolic interactions in non-human primates

Shao-Ming Gao,[1] Li-Ying Lan,[1,2] Li Yang,[1] Tao Chen,[3] and Peng-Fei Fan[1,4,*]
[1]School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, P.R. China
[2]Behavioral Ecology and Sociobiology Unit, German Primate Center, Leibniz Institute for Primate Research, 37077 Göttingen, Germany
[3]School of Ecology, Sun Yat-sen University, Guangzhou 510275, P.R. China
[4]Lead contact
*Correspondence: fanpf@mail.sysu.edu.cn
https://doi.org/10.1016/j.celrep.2025.116477

SUMMARY

Gut microbiota often undergo metabolic cross-feeding and resource competition. However, our understanding of global variations in these interactions and their implications for host health remain elusive. By analyzing a microbial genome catalog from 841 fecal metagenomes across 53 primate species worldwide, we identified key microbiota assigned to two taxa, i.e., Bacillota_A and Pseudomonadota, which well predicted the trade-off of community-level interaction types between metabolic competition and cooperation. Specifically, Bacillota_A species were inherently competitive and amino acid auxotrophic and typically found in anaerobic habitats. In contrast, members of Pseudomonadota were inherently cooperative, siderophore producers, and more abundant in aerobic conditions. Random forest models successfully distinguished unhealthy gut samples from healthy samples through the key competitive and cooperative microbiota, suggesting potential links between community metabolic interactions and host health. Together, this study enhances our mechanistic understanding of microbial interaction dynamism within complex gut ecosystems, offering new targets for understanding host health.

INTRODUCTION

The gut harbors diverse and abundant microbial life akin to an functional organ of their host.[1,2] These taxonomically and functionally differentiated microbial inhabitants typically interact with one another in a variety of ways.[3–5] Generally, the ecological interaction types are defined based on the consequences of the interaction has on the fitness of the co-culturing partners, including both synergistic effects (often as cooperation) and antagonistic effects (predominately referred to as competition).[3,6] These microbial interactions play crucial roles in host metabolism and health.[7–10] However, gut microorganisms do not exist in isolation but form complex assemblies in which myriad microbial species interact with each other at a community level.[5] According to the community assembly and ecological network theory, species interactions govern the gut microbiome dynamics and stability, which are often considered critical for host health.[11,12] Consequently, there is an urgent need in identifying microbial interacting patterns and their underlying rules at the community level in the complex gut ecosystem.

Traditionally, different patterns of microbial interactions are identified using co-cultivation dependent method.[6] However, the vast majority of microbes in the gut have thus far eluded cultivation,[13–15] rendering it formidable to disentangle the nature of species interactions of a community. As a result, previous studies either focused on the interactions of cultured core microbial consortium or considered pairwise co-occurrence as a proxy of interaction,[3,16,17] obscuring our comprehensive understanding of microbial interactions in community. Recent advances in sequencing technologies and metagenomics have enabled the use of genome-scale metabolic models to infer the metabolic competitive and cooperative potential between organisms from the degree of metabolic resource overlap (MRO) and metabolic interaction potential (MIP) between them.[18–20] A higher community MRO indicates a greater overlap between the minimal nutritional requirements of all member species, while a higher MIP suggests more metabolites species can share to decrease their dependency on external resources.[21] Studies have attested to the potential of this method to evaluate the inherently complex microbial metabolic interactions in natural and gut environments.[22–24] Therefore, employing metabolic interaction methods could provide valuable insights into the fundamental rules governing the ecological interactions within the gut microbial communities, thereby facilitating the identification of key microbiome members and streamlining our understanding of complex gut ecosystems.

Previous studies have attested to the roles of many factors in shaping the compositional makeup and functional performance of the gut microbiome.[25–28] For example, the gut microbiome of wild animals is subject to the captivity as a result of human

manipulation of the diet, healthcare, and social interactions.[29] However, while considered across host lineages, phylogeny outperforms other factors (e.g., diet, individual, living environment, and social interactions) in determining the structure of gut microbial community.[30] Therefore, we hypothesize that the gut microbial metabolic interactive patterns are also mainly driven by the host phylogeny (hypothesis 1). Given that functional redundancy and diversification are prevalent in the gut microbiome,[31] the metabolic competition and cooperation for different microorganisms in the community could be polarized.[20] We thus hypothesize that the intrinsically competitive or cooperative key microbiota determines the interaction types of the whole community (hypothesis 2). Finally, the gut microbial interactions are associated with a number of important functions in the host, including the breakdown of complex carbohydrates,[32] defend against pathogenic bacteria,[7] and production or biotransformation of metabolites (i.e., short chain fatty acids, vitamins, and bile acid),[33–35] which are tightly associated with host health. We thus anticipate that shifts in the key competitive and cooperative microbiota that influence the community interactions are associated with changes in host health status (hypothesis 3).

To test these hypotheses, we firstly constructed a comprehensive non-human primate (NHP) gut microbial genome catalog (termed NPG-MG), established with previous gut metagenomic data across diverse clades of NHPs in the globe, as well as our sampling and metagenomic sequencing of feces from two sympatric primate species in China (Table S1). Using this catalog, we sought to uncover the underlying drivers and mechanisms governing the variation in metabolic interactions of the gut microbial community across NHPs. By including an additional dataset comparing the gut microbiome between hosts with distinct health states,[36] we also explore the potential links between the gut microbial metabolic interactions and host health. As a whole, these findings presented in our study could substantially improve our understanding of the complex gut microbial ecosystem and advance the conservation and management of NHPs by leveraging the gut microbial metabolic interactions as indicators for disease diagnosis.

RESULTS

Construction of an NPG-MG catalog

To obtain a comprehensive genome collection for the NHPs gut microbiome, we downloaded 725 metagenomes $\geq$1 Gb from 23 previously published articles, spanning 51 primate species from 22 countries in four continents (Figure 1A; Table S1). Meanwhile, we collected 116 fecal samples of the sympatric western black crested gibbons (*Nomascus concolor*, n = 68) and Indochinese gray langurs (*Trachypithecus crepusculus*, n = 48) in Mt. Wuliang National Nature Reserve, Yunnan, China, for metagenomic sequencing (Figures 1A and S1A; Table S1). Finally, our obtained metagenomic datasets included the major clades (ape, monkey, and lemur) of primate phylogeny, comprising 53 extant primate species (9 families and 31 genera) varied in diet (folivore, frugivore, herbivore, omnivore, carnivore, and gumnivore) and living environment (wild and captive; Figures 1A and 1B; Table S1).

A total of ~27 Tb metagenomic reads data were recruited and used for the genome binning, yielding 15,309 high-quality (completeness >90% and contamination <5%) metagenomic-assembled bins. These bins, in combination with the 1,031 high-quality bins recovered in a previous NHP gut microbiome study,[37] were dereplicated into 3,867 species-level genome bins (SGBs) with an average of 96.3% completeness and 0.72% contamination (Figures 1C and S1B; Table S2). Rarefaction analysis indicated that the number of microbial SGBs was closer to saturating when only considering those with more than two conspecific genomes compared with all SGBs, suggesting that rarer microbial species remain to be discovered (Figure 1C). Nevertheless, SGBs recovered in our study resulted in an over 600% increase in phylogenetic diversity compared to the previous genome catalog of NHP gut microbiome (Figure S1C).

We further examined biases within the datasets across various samples and primate lineages. We selected metagenomes with the largest size within the nine primate families and subsampled them at varying fractions to simulate different sequencing depth. All 3,867 non-redundant SGBs were mapped to both the raw and subsampled metagenomic reads. The number of identified SGBs increased with the subsampling fractions across the nine families (Figure S1D). However, the dominant SGBs (relative abundance >0.1%) in the raw metagenomic samples were successfully recovered in the subsampled datasets, with the exception of Cheirogaleidae that had a microbial richness <10 (Figure 1D). These results suggest that while sequencing depth affects genome recovery efficiency, the dominant microbiota in samples with higher richness could be efficiently recovered in metagenomes across varying sequencing depth. Consequently, we used the dominant SGBs in each sample for our subsequent analyses and excluded 28 samples with a microbial richness <10 from the 841 samples. Furthermore, as most bins were predominantly recovered from specific primate lineages, such as apes and monkeys, we also removed samples with reads mappability <50% to minimize potential biases in the recovered SGBs across different NHPs. Finally, a total of 184 metagenomes (representing 20 species from seven NHPs families) and 3,261 SGBs were retained for subsequent analyses (Table S3).

Primate phylogeny determines the gut microbial metabolic interactions

We next assessed the extent of gut microbial niche overlap and cross-feeding in the 184 communities by constructing genome-scale metabolic models for the identified dominant SGBs of each sample. We used the MRO and MIP scores as proxies to reflect the strength of competitive and cooperative metabolism in a community (See STAR Methods section). We observed significant variations in MRO and MIP across host dietary, living environment, gut morphology, and phylogenetic clades (Figure 2A; Table S3). Specifically, the gut microbiota of folivorous and herbivorous primates exhibited a tendency toward increased competition and cooperation intensity compared with frugivorous primates. Captive primates showed significantly higher strength of cooperation but similar levels of competition relative to wild living primates. Foregut fermenters exhibited significantly higher levels of both competition and cooperation compare to hindgut primates. Monkeys demonstrated the highest levels of MRO and MIP, followed by apes and lemurs (Figure 2A), suggesting potential co-diversification of hosts and gut microbes.

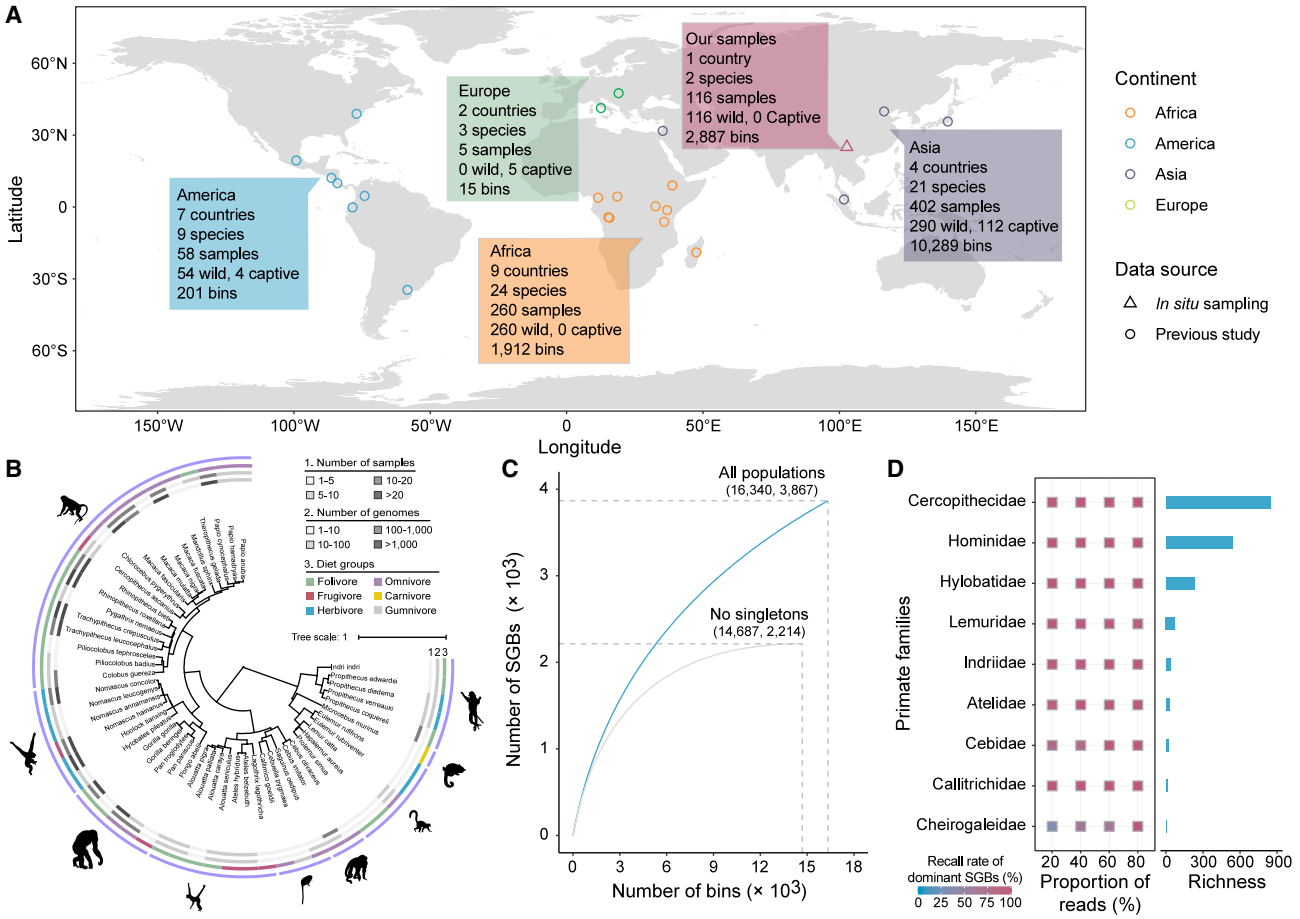


Figure 1. Construction and assessment of the unified NPG-MG catalog

(A) Geographic distribution of the 841 metagenomic samples retrieved per country represented by its capital location.

(B) A consensus phylogenetic tree of the 53 species with gut microbial metagenomes. The colors of the bars in the third inside rings represent the number of metagenomic samples, the number of recovered microbial genomes, and the diet per species. The outer gray strips represent different primate families.

(C) Rarefaction curves of the number of redundant microbial populations detected as a function of the number of genomes analyzed. Curves are depicted for all the populations (blue curve) and after excluding singleton populations (gray curve, represented by only one genome). Numbers inside the parentheses indicate the total number of bins and SGBs.

(D) The recall rate (heatmap) of dominant SGBs in the metagenomes subsampled from nine selected raw metagenomes with the largest size within each primate family. Barplot indicates the number of SGBs detected in the raw metagenomes.

See also Figure S1 and Tables S1 and S2.

Therefore, we performed phylogenetically controlled mixed models (MCMCglmm) to explicitly explore the relationship between the gut microbial interaction strength and host characteristics. The MCMCglmm models converged successfully with the Gelman-Rubin criterion <1.01 (Figure S2). The analysis revealed that primate phylogenetic relationship exerted influence on both gut microbial MRO and MIP levels, as indicated by the significant phylogenetic signal lambda for MCMCglmm models (Figure 2B). We next sought to identify whether interactive types discerned through MRO and MIP scores are phylogenetically related in our samples (Figure 2C). By comparing the interaction indices between each community and random assemblies of the same size, we classified the 184 samples into four interaction types (Figure 2C). Specifically, 20 samples were categorized as both competitive and cooperative (type 1), 16 samples as exclusively cooperative (type 2), 15 samples as neither competitive nor cooperative (type 3), and 133 samples as exclusively competitive (type 4). Further inspection showed that a majority of ape-derived (70.0%) and monkey-derived (75.5%) samples fell into type 4, while most lemur-derived samples (80.0%) were found in type 2 (Figure 2C). However, when the host phylogeny was limited to the species level, only one interactive type was exclusively identified in only eight out of 26 species (Figure S3A), suggesting a weak influence of host phylogeny on interaction types. We further explored the nuanced roles of host phylogeny in determining gut microbial metabolic competition and cooperation. Our analysis revealed that the MRO and MIP scores were significantly and positively correlated with the mean genomic relatedness (average amino acid identity [AAI]) and the richness of dominant SGBs across samples, respectively (Figure 2D).

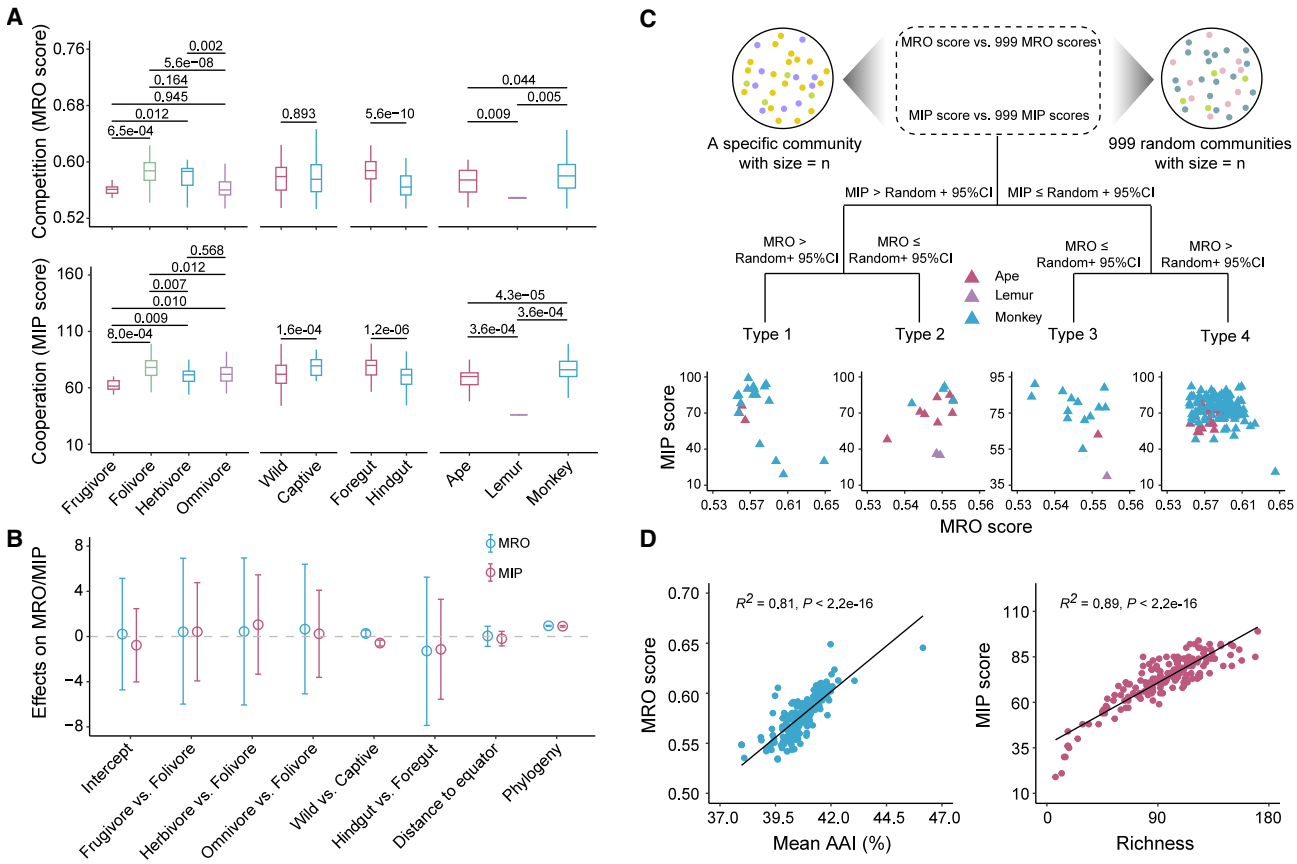


Figure 2. Drivers of gut microbial metabolic interactions in primates

(A) The strength of metabolic competition represented by MRO score (left panel) and metabolic cooperation represented by MIP score (right panel) across different host diets, living environments, gut morphology, and phylogenetic groups. The *p* values (pairwise Wilcoxon test) were adjusted by multiple testing corrections using the Benjamini and Hochberg false discovery rate (FDR) controlling procedure.

(B) Effects of host characteristics on gut microbial MRO and MIP scores. Posterior means (circles) and their 95% CIs (lines) are plotted from phylogenetically controlled MCMCglmm models. The effects of host phylogeny on MRO and MIP scores were calculated as the posterior probability of the phylogenetic signal lambda for MCMCglmm models, i.e., the proportion of the variance components compromised by the phylogenetic variance component.

(C) Identification of samples belonging to the four different interaction types by comparing MRO/MIP scores of specific microbial community with those of random communities (permutations = 999) with the same number of SGBs.

(D) Associations between the richness and average amino acid identity (AAI) and the MRO and MIP scores across the 184 gut microbial communities. Only dominant (relative abundance $\geq$ 0.1%) SGBs are considered in each community. The best-fit lines and Pearson's R^2 values are presented.

See also Figures S2 and S3 and Table S3.

These two gut microbial properties were also influenced by host phylogeny, as indicated by the converged MCMCglmm models (Figures S2 and S3B). Consequently, host phylogeny exerted distinct influences on the gut microbial metabolic competition and cooperation. These findings partially support our first hypothesis.

Key microbiota of *Bacillota_A* and *Pseudomonadota* drive the metabolic interactive types

Having illustrated the gut microbial metabolic interaction types and their driving factors, we further explored the contribution of individual microbial lineages to the variation of interaction types. We defined the contribution of a microbial SGB to the community competition or cooperation as the decrease in MRO or MIP scores when the SGB is deleted from the community. Accordingly, our analysis showed that Bacillota_A, Bacillota_C, and Spirochaeota were the primary contributors to the community MRO, while Pseudomonadota and Spirochaeota predominantly contributed to the MIP among all samples (Figure 3A). Further inspection showed that the proportion of SGBs related to Bacillota_A were significantly higher in competitive samples (type 1 and 4), while that of Pseudomonadota SGBs were significantly higher in cooperative samples (type 1 and 2). Moreover, the proportion of Bacillota_A and Pseudomonadota were significantly correlated with MRO and MIP scores across samples, respectively (Figure S4A), implying the potential key roles of these two phyla in the variation of interaction types.

We further defined an SGB with highest contribution to MRO/MIP in a community as the key SGB that could result in the largest decrease in the community interaction strength.

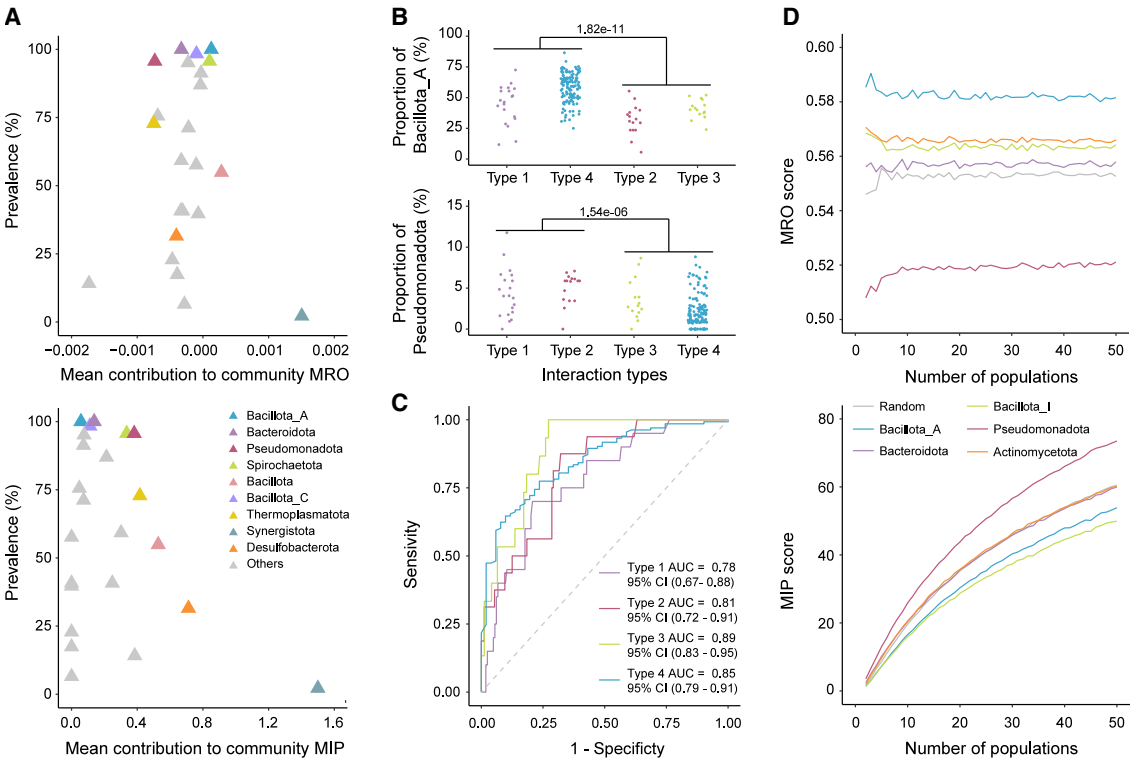


Figure 3. Identification of key gut microbiota contributing to the metabolic interactions

(A) The prevalence and mean contribution of each phylum to MRO (top panel) and MIP (bottom panel) variations. The contribution of an SGB to MRO or MIP of a community is defined as the variation in MRO or MIP when the SGB is deleted from the community. The mean contribution of a phylum to MRO/MIP is then calculated as the mean value of the contribution of SGBs assigned to this phylum in all communities. For clarity, only phyla with the highest prevalence or mean contribution are colored.

(B) Proportion of SGBs assigned to Bacillota_A (top panel) and Pseudomonadota (bottom panel) across samples of the four interaction types. Wilcoxon test was used for comparison of the proportions between competitive (types 1 and 4) or cooperative (types 1 and 2) types and other types.

(C) Based on the presence-absence matrix of the 42 key SGBs assigned to Bacillota_A and Pseudomonadota across the 184 samples, a random-forest classification model with leave-one-out cross-validation was trained to classify different interaction types. The receiver operating characteristic (ROC) curve and the area under the curve (AUC) are shown for each type.

(D) Line plots show the average value of metabolic interaction indices as a function of community size (permutations = 999 for each size). For communities with different phyla, see STAR Methods section and Figure S4B for detailed simulation steps. For the random communities, the SGBs were randomly selected from the 3,867 SGBs.

See also Figure S4 and Table S4.

Accordingly, 29 key SGBs assigned to Bacillota_A, and 13 key SGBs affiliated with Pseudomonadota were identified for the community MRO and MIP, respectively (Figure S4B). The 29 key Bacillota_A SGBs were mainly affiliated with Lachnospiraceae, Ruminococcaceae, and Acutalibacteraceae, such as *Roseburia hominis*, *Hominimerdicola* spp., and *UBA737* spp., and the 13 key Pseudomonadota SGBs were mainly affiliated with Burholderiaceae, CAG-239, Succinivibrionaceae, and Enterobacteriaceae, such as *Sutterella* spp., *Ralstonia* spp., *Succinivibrio* spp., and *Escherichia coli* (Table S4). We estimated the presence-absence matrix of these 42 key SGBs across the 184 samples to develop a random forest classifier to distinguish different interaction types. Receiver operating characteristic (ROC) curve analysis demonstrated a moderate-to-high diagnostic capability, with an area under the ROC curve (AUC) varied for different types, ascertained through leave-one-out cross-validation (Figure 3C).

To investigate whether the contributions of these two phyla were stemmed from their intrinsic nature of metabolic interactions, we performed simulations for different sized communities comprising a random number of SGBs belonging to specific phyla (Figure S4C). We calculated MRO and MIP scores for the simulated communities of the five phyla (Bacillota_A, Bacteroidota, Bacillota_I, Pseudomonadota, and Actinomycetota) with the highest richness (87.6% of the total; Figure S4D). Our analysis revealed that the MRO scores remained consistent, while the MIP scores increased with community size across different phyla (Figure 3D). Meanwhile, the simulation results showed that communities with Bacillota_A were highly competitive, while those with Pseudomonadota displayed the highest degree of cooperation (Figure 3D). Altogether, these results provide evidence for our second hypothesis that the key competitive and cooperative species drive the variation in metabolic interaction types of the gut microbial community.

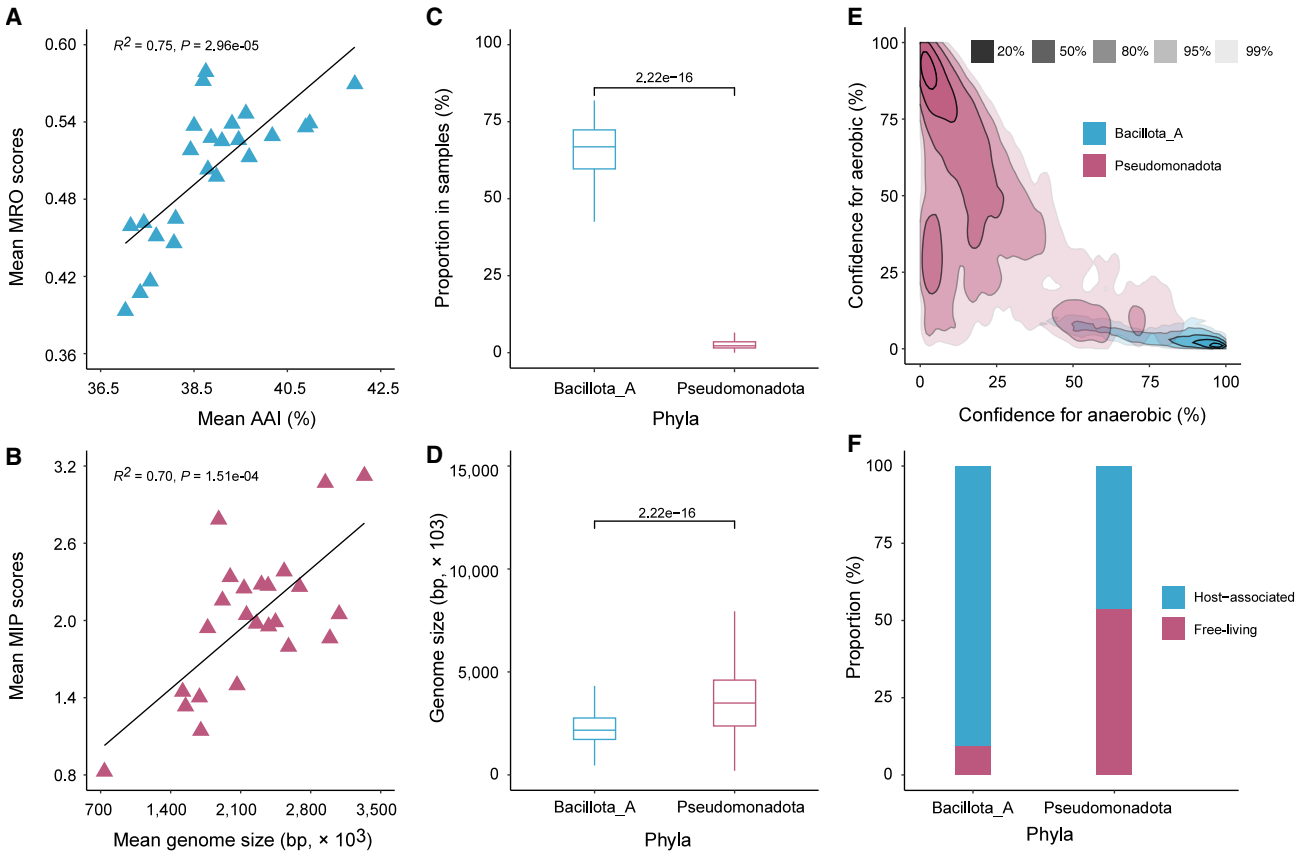


Figure 4. Genomic and ecological disparities between Bacillota_A and Pseudomonadota

(A and B) The associations between the genomic features, i.e., mean genomic relatedness (A) and mean genome size (B), and the metabolic interaction scores across different phyla. For each SGB, the average amino acid identity (AAI) values or metabolic interaction scores between this SGB and all other 3,866 SGBs are estimated and averaged. The mean genomic relatedness or metabolic interaction scores for a specific phylum is thus determined as the average AAI value or MRO/MIP scores of SGBs assigned to that phylum. The best-fit lines and Pearson's R^2 values are presented.

(C) The proportion of SGBs assigned to Bacillota_A and Pseudomonadota across the 184 samples. Wilcoxon test was used for comparison of the proportion.

(D) The genome size of SGBs in the NPG-MG catalog and representatives in the GTDB r220 database affiliated with Bacillota_A and Pseudomonadota. Wilcoxon test was used for comparison of the proportion.

(E) Proportion of Bacillota_A and Pseudomonadota genomes isolated from host-associated or free-living environments in the GTDB r220 database.

(F) 2D kernel density plot of the distribution of confidence values of the anaerobe (*x* axis) and aerobe (*y* axis) models of genomes in the NPG-MG catalog and the GTDB r220 database affiliated with Bacillota_A and Pseudomonadota. The color gradient is proportional to the expected probability levels.

See also Figures S5 and S6.

Genomic and ecological evidence of Bacillota_A and Pseudomonadota for their polarized metabolic interactions

We next investigated the genomic and ecological adaptations of Bacillota_A and Pseudomonadota to their polarized metabolic interactions in the gut. Given that the MRO and MIP scores were defined based on the input or output compounds of microbial genomes, we initially examined the associations between the scores and genomic features. The mean MRO score exhibited a positive correlation (R^2 = 0.75, p = 2.96e-05) with the mean genomic relatedness (AAI) across phyla (Figure 4A, top panel). Meanwhile, the mean MIP score showed a positive correlation (R^2 = 0.70, p = 1.51e-04) with the average genome length across phyla (Figure 4A, bottom panel). Specifically, Bacillota_A SGBs displayed significantly higher genomic relatedness with other genomes (Figure S5A), which may be attributed to their prevalent distribution in the gut (Figure 4C). In contrast, the significantly larger genome size of Pseudomonadota reflects greater metabolic capacity (Figure 4D), as evidence by the larger number of synthesized amino acids and utilized carbon sources in their genomes (Figures S5B and S5C).

Further analyses revealed that genome size was more strongly correlated with the number of output compounds than with the number of input compounds (Figures S6A and S6B). Thus, we conducted investigation into the exogenously absorbed substrates by Bacillota_A SGBs and the endogenously secreted substrates by Pseudomonadota SGBs. Amino acids were found to be the most frequently required compounds for the Bacillota_A SGBs (Figure S6C), mirrored that this lineage displayed lower capacity for amino acid biosynthesis (Figure S5B). The Pseudomonadota SGBs exhibited a wider distribution in their contributed compounds to the community (Figure S6D). Notably,

siderophores (i.e., Salmocheline-S4 and Enterochelin) produced by Pseudomonadota SGBs emerged as prominent public goods in the community (Figure S6D).

To gain insights into the potential link between genomic and ecological adaptations of these two phyla, we further investigated their natural distribution. We firstly assessed the oxygen requirements of Bacillota_A and Pseudomonadota genomes. The results indicated that Bacillota_A species were predominantly anaerobic, whereas Pseudomonadota species were primarily aerobic (Figure 4E). We then examined the habitat types for their isolated genomes in Genome Taxonomy Database (GTDB) r220. The results showed that 90.8% of the genomes affiliated with Bacillota_A originated from host-associated environments, while 53.7% of the Pseudomonadota genomes were found in free-living conditions (Figure 4F). Concurrently, community composition data obtained from the Earth Microbiome Project (EMP) also indicated that Bacillota_A was more abundant in host-associated environments (mean relative abundance = 16.4%) compared to free-living ones (1.56%), whereas Pseudomonadota exhibited significantly higher relative abundance in free-living environments (37.1%) than host-associated environments (36.0%) (Figure S5D). Collectively, these results suggest distinct oxygen requirements and habitat preferences of these two phyla: competitive Bacillota_A characterized by a smaller genome size tends to dominate in anaerobic and host-associated environments, while cooperative Pseudomonadota with a larger genome size is more abundant or prevalent in aerobic and free-living conditions.

Case study shows potential health implications of gut microbial metabolic interactions

To further elucidate the potential associations between gut microbial metabolic interactions and host health, we conducted an analysis of sequenced amplicon data recruited from a previous comparative study focusing on the gut microbiome of healthy and unhealthy NHPs.[36] We firstly mapped the 16S rRNA gene sequences to their closest genomes in the NPG-MG and the GTDB r220 representatives using a threshold of 99% identity and 95% coverage. Consequently, 52.3% of amplicon sequence variants (ASVs) were successfully mapped to 1,257 assembled genomes (Table S5), among which 81 were from NPG-MG and 1,176 were from the GTDB database. Notably, the mapped ASVs exhibited significantly higher prevalence and abundance compared to unmapped ASVs (Figure S7A), indicating that unmapped ASVs are less prevalent and abundant within the community.

We then used these mapped genomes to analyze the community composition and metabolic interactions. The gut microbiome of both the healthy and unhealthy NHPs were dominated with Bacteroidota, Bacillota_A, Pseudomonadota, and Bacillota (Figure 5A). In particular, our analyses detected that Bacillota_A was significantly more abundant in healthy individuals, whereas Pseudomonadota was significantly enriched in unhealthy individuals (Figure S7B). Additionally, from healthy to unhealthy primates, we exclusively observed an increase in the proportion of Pseudomonadota, while no significant alteration in the proportion of Bacillota_A was detected (Figure 5B). Concomitantly, significant increase in the strength of gut microbial metabolic cooperation was observed in unhealthy guts compared to healthy ones (Figure 5B).

To identify the key species that determined the community metabolic interactions, we assessed the contribution of each species to the MRO/MIP in each community. Consistently, Bacillota_A and Pseudomonadota species contributed most to the metabolic competition and cooperation across communities, respectively (Figure 5C). Thus, we identified 42 Bacillota_A species and 25 Pseudomonadota species as the key species that contributed mostly to MRO/MIP in their samples (Table S6). Among the 67 key species, only three were mapped to genomes in the NPG-MG catalog, while the others were mapped to the GTDB representatives (Table S6). The 42 key Bacillota_A SGBs were mainly affiliated with Lachnospiraceae, Ruminococcaceae, and Oscillospiraceae, such as *Roseburia hominis*, *Ruminococcus* spp, and *CAG-170* spp., and the 25 key Pseudomonadota SGBs were mainly affiliated with Burholderiaceae, Pasteurellaceae, Moraxellaceae, and Enterobacteriaceae, such as *Comamonas* spp., *Aggregatibacter actinomycetemcomitans*, *Acinetobacter* spp., and *E. coli*. Using the presence-absence matrix of these key species across samples, we trained a random forest classifier to distinguish unhealthy from healthy primates. These key species showed excellent diagnostic power in the datasets, achieving an AUC value of 0.99 (Figure 5D).

We finally sought to understand the underlying mechanisms of the potential links between community metabolic interactions and host health status. We performed functional annotation to identify virulent factor genes (VFs) and antibiotic resistance genes (ARGs) in the genomes. Among the 1,257 representative genomes mapped to ASVs, Pseudomonadota species harbored the highest average numbers of VFs and ARGs, while Bacillota_A had relatively lower VFs and ARGs (Figure 5E). This pattern remained consistent when including Bacillota_A and Pseudomonadota genomes recovered in our NPG-MG catalog (Figure S5C). Collectively, these results provide evidence for our third hypothesis that the metabolic interactions of gut microbial community could be a potential indicator of host health in NHPs.

DISCUSSION

We compiled a unified NPG-MG catalog of primate gut microbiomes, encompassing 3,867 microbial species distributed across 53 global primate animal species. While this catalog significantly expanded our knowledge about the NHPs gut microbial phylogenetic diversity compared to previous study,[37] we observed that most of our samples and microbial populations originated from catarrhine monkeys and apes. Therefore, continued efforts to investigate the gut microbiome of platyrrhine monkeys and lemurs will further improve the resolution of our catalog. Nevertheless, the NPG-MG represents the most extensive catalog to date capturing the breadth of gut microbial genomic diversity across extant primates and, thus, will serve as a comprehensive reference for detailed exploration of metabolic interactions in gut. Additionally, the microbiota of NHPs has received less attention compared to that of the humans.[37] However, over two-third of primate species are threatened with extinction.[38] Given that mapping the microbiome of wildlife

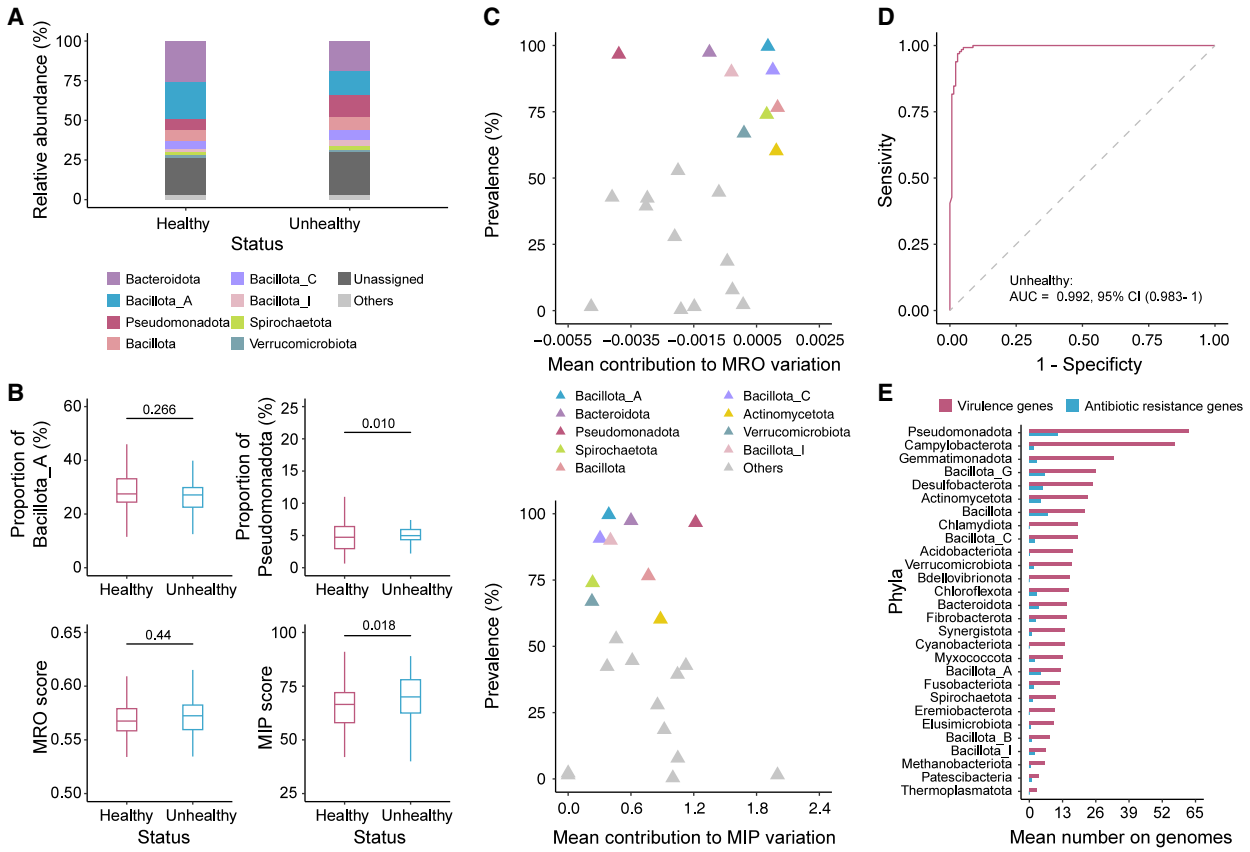


Figure 5. Case study of variations in gut microbial metabolic interactions of NHPs with different health status

(A) The relative abundance of gut microbial communities in samples of healthy and unhealthy NHPs.

(B) Proportion of Bacillota_A and Pseudomonadota genomes (top panel), and the strength of metabolic competition represented by MRO score and the strength of metabolic cooperation represented by MIP score (bottom panel) in samples of healthy and unhealthy NHPs. Wilcoxon test was used for the comparisons.

(C) The prevalence and mean contribution of each phylum to MRO (top panel) and MIP (bottom panel) variations. For clarity, only phyla with the highest prevalence or mean contribution are colored.

(D) Based on the presence-absence matrix of the 67 key SGBs assigned to Bacillota_A and Pseudomonadota across the samples, a random-forest classification model with leave-one-out cross validation was trained to classify healthy and unhealthy samples. The ROC curve and the AUC are shown.

(E) The average number of identified ARGs and VFs on the mapped genomes of different phyla.

See also Figure S7 and Tables S5 and S6.

could guide conservation efforts,[39] our studies would have an important significance for science-informed conservation. Furthermore, recent studies have suggested the co-diversification between the gut microbiome and primates, highlighting the widespread extinction of ancestral microbes from the human gut.[40] Therefore, the NPG-MG catalog would not only help in elucidating co-evolutionary trajectories shaping the current structure of the human microbiome but also could provide targets for preserving gut microbiota in the endangered primates.

Our study unveiled that host phylogeny plays a crucial role in shaping the gut microbial metabolic interactions in primates. The concept of phylosymbiosis has been developed as a general principle for understanding the eco-evolutionary processes governing the gut microbiome structure.[30,40] In light of this perspective, our conclusion is intuitive, yet it has not been previously reported, possibly because of substantial emphasis on key enterotype taxa instead of their biotic interactions in the gut.[41] While a previous study using high-order co-occurrence uncovered stark partitioning between competing and cooperating groups,[20] our results indicated that the strength of metabolic competition and cooperation do not always exhibit a polarized, antagonistic relationship in natural communities. This result is reasonable as the competitive and cooperative microbial groups could be concurrently present in a single natural community.[42,43] While the gut microbial metabolic interactions could be categorized into different types, analogous to the concept of enterotype,[44] the gut microbial metabolic competition and cooperation were driven by different factors, thereby weakening the overall impact of host phylogeny on community interaction types derived simultaneously from the two interaction scores. Nonetheless, our study elucidated that a majority of samples belong to type 4, implying that most gut microbial communities are metabolic competitive. This result potentiates the fact that the gut microbiome is relatively stable as competition can promote community stability.[11] However, future research should take microbial abundance and other microbial interactions types

(e.g., phage predation, and toxic or antibiotic effects) into account, as these aspects are critical to the microbial interaction types in the gut.[45,46]

Our analyses identified that Bacillota_A and Pseudomonadota contributed most to the gut microbial metabolic interactions in NHPs. More importantly, when compared with other phyla and random communities, simulated communities, including SGBs of Bacillota_A and Pseudomonadota, showed the highest competition risk and cooperation potential, respectively, attesting to their inherent property of metabolic interactions. Members of Bacillota_A usually have significantly higher proportions of amino acid auxotrophs than most other phyla,[47] endorsing our results that Bacillota_A harbor less amino acid biosynthesis pathways and would frequently absorb amino acids exogenously. Meanwhile, considering that Bacillota_A species contribute mostly to the diversity in the gut microbial community of NHPs, our results are consistent with a previous study, which claimed that the gut microbiome with higher diversity is characterized with increased auxotrophy frequency.[48]

Our study also provided genomic and ecological evidences that potentiate MIP. Genomically, the auxotrophic Bacillota_A species would compete for the depleted amino acids in the colon where the majority of gut microbiota resides,[48] as diet-derived amino acids are mostly absorbed by the host in the upper gastrointestinal tract.[49] However, Pseudomonadota displays larger genome size and secrets siderophores as public goods that scavenge iron from environmental stocks, leading to potential cooperative interactions between microorganisms.[50] Ecologically, due to the polarization of oxygen requirements between the two phyla, Bacillota_A species are specifically found in host-associated environments that are usually anaerobic, e.g., the gut, whereas Pseudomonadota species tend to be more prevalent and abundant in free-living conditions that are commonly aerobic. Since free-living communities typically feature a larger genome size, enlarged genome size variability, and higher proportions of prototrophs,[51] reduced competition and increased cooperation are expected in such habitat for all of these reasons. Notably, our findings contrast results from a previous work on co-occurring microbial species where competitive groups are dominantly found in free-living environments.[20] Given that co-occurring species exist as only a part of the whole community, it is likely that metabolic interactions with other species play a role in the community.

Our study has uncovered key gut microbiota that underpin our machine-learning models to successfully differentiate not only among various interaction types but also between different host healthy statuses. The intricate interactions within gut microbial communities play crucial roles in shaping host functions.[3,4] Therefore, identifying key microbiota with predictive capabilities for community-level metabolic interactions would advance our understanding of the complex gut microbiome system and enable anticipating how interaction dynamics might shift in response to external perturbations. Notably, the identified key Bacillota_A species (e.g., *Roseburia hominis*, *Hominimerdicola* spp., *Ruminococcus* spp., *UBA737* spp., and *CAG-170* spp.) are prevalent and abundant in the gut. These species are specified in the degradation of dietary fiber and production of short-chain fatty acids (SCFAs) and showed reduced abundance levels in unhealthy hosts,[52–56] indicating a fundamental role of these species in the healthy gut. In contrast, while the cooperative Pseudomonadota are typically rarer in the gut, members of this phylum are usually opportunistically pathogenic,[57] as evidenced by the highest number of VFs and ARGs in genomes. Meanwhile, the identified key Pesudomonadota species (e.g., *Sutterella* spp., *Ralstonia* spp., *Succinivibrio* spp., *Comamonas* spp., *A. actinomycetemcomitans*, *Acinetobacter* spp., and *E. coli*.) have shown positive correlations with diseases,[58–64] suggesting a detrimental role of these species in the gut. Consequently, the balance between the gut microbial metabolic competitive and cooperative types might also imply the host health status. This is supported by our case analyses, which revealed significantly higher metabolic cooperation levels in unhealthy NHP hosts. It has been suggested that reduced diversity and cooperation-derived dependencies foster instability in the gut ecosystem.[11] While no significant differences in diversity were observed between healthy and unhealthy hosts,[36] our results indicate that the microbial community metabolic strategy might serve as an indicator of host health, suggesting the independent roles of microbial diversity and interactions in determining community stability.[65] Moreover, our classification and prediction models demonstrated the robustness of using key microbiota contributing to the metabolic interactions to accurately predict the healthy status of NHPs, providing new strategies to promote their management and conservation. Notably, only a small proportion of ASVs could be mapped to genomes in the NPG-MG catalog, as the catalog included just two of the five NHPs species analyzed in the case study.[36] This result highlights the urgent need to construct a more comprehensive NHP gut microbiome catalog derived from individuals with diverse health states or diseases, in order to develop a species-independent approach for improving NHPs' health prediction.

The astounding abundance and diversity of microbes colonizing the gut are testament to their collective metabolic capabilities relevant to host fitness and health. Our findings indicate that host evolutionary history profoundly determines the variation of gut microbial metabolic interactions, primarily manifested in the variation of the key microbial taxa, i.e., Bacillota_A and Pseudomonadota. Moreover, we show that the inherent competition and cooperation capacity of these two phyla might be a result of genomic and ecological adaptation. Finally, we reveal that the variation of gut microbial metabolic interactions is accompanied with NHPs' healthy status, suggesting potential therapeutic implications of metabolic interactions. Overall, our comprehensive investigation of gut microbial metabolic interactions in NHP species highlights that gut microbial metabolic interaction manifests itself as a multi-type community, where individual gut microbiota congregates around competitive and cooperative taxa. Our study provides new insights into the way we think about the complex microbial ecosystem in general. On the one hand, the microbial interactions play a crucial role in community stability and functionality. If the recurring pattern that gut microbial interactions are governed by specific relationships between particular group of microorganisms is confirmed, we can streamline this complex system by focusing on these key interactions, thereby facilitating future rational manipulations. On the other hand, if such multi-type communities are prevalent across

various ecosystems, uncovering diverse key taxa contributing to the interaction types and their underlying selection forces would promote our understanding of the origins and evolutionary trajectories of microorganisms on the Earth.

Limitations of the study

In this study, we demonstrated that key gut microbiota assigned to Bacillota_A and Pseudomonadota drives the variation in community-level metabolic interactions across NHPs. It is noteworthy that our current analysis may not fully account for the influence of species abundance and other microbial interactions types (e.g., phage predation, and toxic or antibiotic effects) on the intensity of microbial interactions. Future studies employing either microcosm experiments or *in silico* simulations are needed to address this limitation, thereby enhancing our comprehensive understanding of complex gut microbial interactions. Additionally, while we observed associations between gut microbial metabolic interactions and host health status, the underlying mechanisms of microbial functional roles in maintaining gut homeostasis remain to be elucidated at a more granular level.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and fulfilled by the lead contact, Peng-Fei Fan (fanpf@mail.sysu.edu.cn).

Materials availability

Materials generated in this paper are available from the lead contact upon request.

Data and code availability

- Data: Sequencing data from this study have been deposited in the NCBI BioProject database. Accession code (PRJNA1158499) is also listed in the key resources table. Metabolic models for the 3,867 SGBs have been deposited at the Zenodo page (https://doi.org/10.5281/zenodo.17140653).
- Code: All original code and relevant data required to reanalyze the data reported in this paper have been deposited at the Zenodo page (https://doi.org/10.5281/zenodo.17139167).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We would like to thank the State Key Laboratory of Biocontrol (Sun Yat-sen University) for assistance with experiments and the Dazhaizi Gibbon Research Station in Mt. Wuliang National Nature Reserve for field work. This work was funded by the Ministry of Science and Technology of China: National Key Research and Development Program (grant no. 2022YFF1301500 to P.-F. F.), the National Natural Science Foundation of China (no. 32171485 to P.-F. F. and no. 32201269 to S.-M.G.), the Guangdong Basic and Applied Basic Research Foundation (no. 2023A1515010695 to S.-M.G.), and Fundamental Research Funds for the Central Universities, Sun Yat-sen University (no. 23lgzy002 to P.-F.F.).

AUTHOR CONTRIBUTIONS

Conceptualization, P.-F.F. and S.-M.G.; methodology, P.-F.F. and S.-M.G.; sampling, L.-Y.L.; experiments, L.-Y.L. and S.-M.G.; formal analysis, S.-M.G.; writing – original draft, S.-M.G.; writing – review and editing, P.-F.F., T.C., L.-Y.L., and L.Y.; supervision, P.-F.F.; and funding acquisition, P.-F.F. and S.-M.G.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2025.116477.

Received: October 17, 2024
Revised: May 22, 2025
Accepted: October 3, 2025

REFERENCES

1. Fan, Y., and Pedersen, O. (2021). Gut microbiota in human metabolic health and disease. Nat. Rev. Microbiol. *19*, 55–71. https://doi.org/10.1038/s41579-020-0433-9.
2. O'Hara, A.M., and Shanahan, F. (2006). The gut flora as a forgotten organ. EMBO Rep. *7*, 688–693. https://doi.org/10.1038/sj.embor.7400731.
3. Coyte, K.Z., and Rakoff-Nahoum, S. (2019). Understanding competition and cooperation within the mammalian gut microbiome. Curr. Biol. *29*, R538–R544. https://doi.org/10.1016/j.cub.2019.04.017.
4. Culp, E.J., and Goodman, A.L. (2023). Cross-feeding in the gut microbiome: Ecology and mechanisms. Cell Host Microbe *31*, 485–499. https://doi.org/10.1016/j.chom.2023.03.016.
5. Zengler, K., and Zaramela, L.S. (2018). The social network of microorganisms - how auxotrophies shape complex communities. Nat. Rev. Microbiol. *16*, 383–390. https://doi.org/10.1038/s41579-018-0004-5.
6. Mitri, S., and Foster, K.R. (2013). The genotypic view of social interactions in microbial communities. Annu. Rev. Genet. *47*, 247–273. https://doi.org/10.1146/annurev-genet-111212-133307.
7. Caballero-Flores, G., Pickard, J.M., and Núñez, G. (2023). Microbiota-mediated colonization resistance: mechanisms and regulation. Nat. Rev. Microbiol. *21*, 347–360. https://doi.org/10.1038/s41579-022-00833-7.
8. Henriques, S.F., Dhakan, D.B., Serra, L., Francisco, A.P., Carvalho-Santos, Z., Baltazar, C., Elias, A.P., Anjos, M., Zhang, T., Maddocks, O.D.K., and Ribeiro, C. (2020). Metabolic cross-feeding in imbalanced diets allows gut microbes to improve reproduction and alter host behaviour. Nat. Commun. *11*, 4236. https://doi.org/10.1038/s41467-020-18049-9.
9. Koppel, N., Maini Rekdal, V., and Balskus, E.P. (2017). Chemical transformation of xenobiotics by the human gut microbiota. Science *356*, eaag2770. https://doi.org/10.1126/science.aag2770.

10. Rakoff-Nahoum, S., Foster, K.R., and Comstock, L.E. (2016). The evolution of cooperation within the gut microbiota. Nature *533*, 255–259. https://doi.org/10.1038/nature17626.

11. Coyte, K.Z., Schluter, J., and Foster, K.R. (2015). The ecology of the microbiome: Networks, competition, and stability. Science *350*, 663–666. https://doi.org/10.1126/science.aad2602.

12. HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M., and Mayfield, M.M. (2012). Rethinking community assembly through the lens of coexistence theory. Annu. Rev. Ecol. Evol. Syst. *43*, 227–248. https://doi.org/10.1146/annurev-ecolsys-110411-160411.

13. Liu, C., Du, M.X., Abuduaini, R., Yu, H.Y., Li, D.H., Wang, Y.J., Zhou, N., Jiang, M.Z., Niu, P.X., Han, S.S., et al. (2021). Enlightening the taxonomy darkness of human gut microbiomes with a cultured biobank. Microbiome *9*, 119. https://doi.org/10.1186/s40168-021-01064-3.

14. Liu, S., Moon, C.D., Zheng, N., Huws, S., Zhao, S., and Wang, J. (2022). Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. Microbiome *10*, 76. https://doi.org/10.1186/s40168-022-01272-5.

15. Nayfach, S., Shi, Z.J., Seshadri, R., Pollard, K.S., and Kyrpides, N.C. (2019). New insights from uncultivated genomes of the global human gut microbiome. Nature *568*, 505–510. https://doi.org/10.1038/s41586-019-1058-x.

16. Kokou, F., Sasson, G., Friedman, J., Eyal, S., Ovadia, O., Harpaz, S., Cnaani, A., and Mizrahi, I. (2019). Core gut microbial communities are maintained by beneficial interactions and strain variability in fish. Nat. Microbiol. *4*, 2456–2465. https://doi.org/10.1038/s41564-019-0560-0.

17. Palmer, J.D., and Foster, K.R. (2022). Bacterial species rarely work together. Science *376*, 581–582. https://doi.org/10.1126/science.abn5093.

18. Gu, C., Kim, G.B., Kim, W.J., Kim, H.U., and Lee, S.Y. (2019). Current status and applications of genome-scale metabolic models. Genome Biol. *20*, 121. https://doi.org/10.1186/s13059-019-1730-3.

19. Lieven, C., Beber, M.E., Olivier, B.G., Bergmann, F.T., Ataman, M., Babaei, P., Bartell, J.A., Blank, L.M., Chauhan, S., Correia, K., et al. (2020). MEMOTE for standardized genome-scale metabolic model testing. Nat. Biotechnol. *38*, 272–276. https://doi.org/10.1038/s41587-020-0446-y.

20. Machado, D., Maistrenko, O.M., Andrejev, S., Kim, Y., Bork, P., Patil, K.R., and Patil, K.R. (2021). Polarization of microbial communities between competitive and cooperative metabolism. Nat. Ecol. Evol. *5*, 195–203. https://doi.org/10.1038/s41559-020-01353-4.

21. Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D.R., Bork, P., and Patil, K.R. (2015). Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc. Natl. Acad. Sci. USA *112*, 6449–6454. https://doi.org/10.1073/pnas.1421834112.

22. Du, H., Pan, J., Zou, D., Huang, Y., Liu, Y., and Li, M. (2022). Microbial active functional modules derived from network analysis and metabolic interactions decipher the complex microbiome assembly in mangrove sediments. Microbiome *10*, 224. https://doi.org/10.1186/s40168-022-01421-w.

23. Heinken, A., Hertel, J., Acharya, G., Ravcheev, D.A., Nyga, M., Okpala, O.E., Hogan, M., Magnúsdóttir, S., Martinelli, F., Nap, B., et al. (2023). Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine. Nat. Biotechnol. *41*, 1320–1331. https://doi.org/10.1038/s41587-022-01628-0.

24. Sung, J., Kim, S., Cabatbat, J.J.T., Jang, S., Jin, Y.S., Jung, G.Y., Chia, N., and Kim, P.J. (2017). Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis. Nat. Commun. *8*, 15393. https://doi.org/10.1038/ncomms15393.

25. Kim, P.S., Shin, N.R., Lee, J.B., Kim, M.S., Whon, T.W., Hyun, D.W., Yun, J.H., Jung, M.J., Kim, J.Y., and Bae, J.W. (2021). Host habitat is the major determinant of the gut microbiome of fish. Microbiome *9*, 166. https://doi.org/10.1186/s40168-021-01113-x.

26. Olm, M.R., Dahan, D., Carter, M.M., Merrill, B.D., Yu, F.B., Jain, S., Meng, X., Tripathi, S., Wastyk, H., Neff, N., et al. (2022). Robust variation in infant gut microbiome assembly across a spectrum of lifestyles. Science *376*, 1220–1223. https://doi.org/10.1126/science.abj2972.

27. Orkin, J.D., Campos, F.A., Myers, M.S., Cheves Hernandez, S.E., Guadamuz, A., and Melin, A.D. (2019). Seasonality of the gut microbiota of free-ranging white-faced capuchins in a tropical dry forest. ISME J. *13*, 183–196. https://doi.org/10.1038/s41396-018-0256-0.

28. Youngblut, N.D., Reischer, G.H., Walters, W., Schuster, N., Walzer, C., Stalder, G., Ley, R.E., and Farnleitner, A.H. (2019). Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. Nat. Commun. *10*, 2200. https://doi.org/10.1038/s41467-019-10191-3.

29. McKenzie, V.J., Song, S.J., Delsuc, F., Prest, T.L., Oliverio, A.M., Korpita, T.M., Alexiev, A., Amato, K.R., Metcalf, J.L., Kowalewski, M., et al. (2017). The Effects of Captivity on the Mammalian Gut Microbiome. Integr. Comp. Biol. *57*, 690–704. https://doi.org/10.1093/icb/icx090.

30. Amato, K.R., Sanders, J.G., Song, S.J., Nute, M., Metcalf, J.L., Thompson, L.R., Morton, J.T., Amir, A., McKenzie, V.J., Humphrey, G., et al. (2019). Evolutionary trends in host physiology outweigh dietary niche in structuring primate gut microbiomes. ISME J. *13*, 576–587. https://doi.org/10.1038/s41396-018-0175-0.

31. Vieira-Silva, S., Falony, G., Darzi, Y., Lima-Mendez, G., Garcia Yunta, R., Okuda, S., Vandeputte, D., Valles-Colomer, M., Hildebrand, F., Chaffron, S., and Raes, J. (2016). Species-function relationships shape ecological properties of the human gut microbiome. Nat. Microbiol. *1*, 16088. https://doi.org/10.1038/nmicrobiol.2016.88.

32. Ye, S., Shah, B.R., Li, J., Liang, H., Zhan, F., Geng, F., and Li, B. (2022). A critical review on interplay between dietary fibers and gut microbiota. Trends Food Sci. Technol. *124*, 237–249. https://doi.org/10.1016/j.tifs.2022.04.010.

33. Mukhopadhya, I., and Louis, P. (2025). Gut microbiota-derived short-chain fatty acids and their role in human health and disease. Nat. Rev. Microbiol. *23*, 635–651. https://doi.org/10.1038/s41579-025-01183-w.

34. Collins, S.L., Stine, J.G., Bisanz, J.E., Okafor, C.D., and Patterson, A.D. (2023). Bile acids and the gut microbiota: metabolic interactions and impacts on disease. Nat. Rev. Microbiol. *21*, 236–247. https://doi.org/10.1038/s41579-022-00805-x.

35. Das, P., Babaei, P., and Nielsen, J. (2019). Metagenomic analysis of microbe-mediated vitamin metabolism in the human gut microbiome. BMC Genom. *20*, 208. https://doi.org/10.1186/s12864-019-5591-7.

36. Amato, K.R., Metcalf, J.L., Song, S.J., Hale, V.L., Clayton, J., Ackermann, G., Humphrey, G., Niu, K., Cui, D., Zhao, H., et al. (2016). Using the gut microbiota as a novel tool for examining colobine primate GI health. Glob. Ecol. Conserv. *7*, 225–237. https://doi.org/10.1016/j.gecco.2016.06.004.

37. Manara, S., Asnicar, F., Beghini, F., Bazzani, D., Cumbo, F., Zolfo, M., Nigro, E., Karcher, N., Manghi, P., Metzger, M.I., et al. (2019). Microbial genomes from non-human primate gut metagenomes expand the primate-associated bacterial tree of life with over 1000 novel species. Genome Biol. *20*, 299. https://doi.org/10.1186/s13059-019-1923-9.

38. Chen, T., Garber, P.A., Zhang, L., Yang, L., and Fan, P. (2023). The pattern and drivers of taxonomic bias in global primate research. Glob. Ecol. Conserv. *46*, e02599. https://doi.org/10.1016/j.gecco.2023.e02599.

39. Wei, F., Wu, Q., Hu, Y., Huang, G., Nie, Y., and Yan, L. (2019). Conservation metagenomics: a new branch of conservation biology. Sci. China Life Sci. *62*, 168–178. https://doi.org/10.1007/s11427-018-9423-3.

40. Sanders, J.G., Sprockett, D.D., Li, Y., Mjungu, D., Lonsdorf, E.V., Ndjango, J.B.N., Georgiev, A.V., Hart, J.A., Sanz, C.M., Morgan, D.B., et al. (2023). Widespread extinctions of co-diversified primate gut bacterial symbionts from humans. Nat. Microbiol. *8*, 1039–1050. https://doi.org/10.1038/s41564-023-01388-w.

41. Larzul, C., Estellé, J., Borey, M., Blanc, F., Lemonnier, G., Billon, Y., Thiam, M.G., Quinquis, B., Galleron, N., Jardet, D., et al. (2024). Driving gut

microbiota enterotypes through host genetics. Microbiome *12*, 116. https://doi.org/10.1186/s40168-024-01827-8.

42. Dai, T., Wen, D., Bates, C.T., Wu, L., Guo, X., Liu, S., Su, Y., Lei, J., Zhou, J., and Yang, Y. (2022). Nutrient supply controls the linkage between species abundance and ecological interactions in marine bacterial communities. Nat. Commun. *13*, 175. https://doi.org/10.1038/s41467-021-27857-6.
43. Liu, H., Liao, C., Wu, L., Tang, J., Chen, J., Lei, C., Zheng, L., Zhang, C., Liu, Y.Y., Xavier, J., and Dai, L. (2022). Ecological dynamics of the gut microbiome in response to dietary fiber. ISME J. *16*, 2040–2055. https://doi.org/10.1038/s41396-022-01253-4.
44. Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.M., et al. (2011). Enterotypes of the human gut microbiome. Nature *473*, 174–180. https://doi.org/10.1038/nature09944.
45. Shkoporov, A.N., Turkington, C.J., and Hill, C. (2022). Mutualistic interplay between bacteriophages and bacteria in the human gut. Nat. Rev. Microbiol. *20*, 737–749. https://doi.org/10.1038/s41579-022-00755-4.
46. Weiland-Bräuer, N. (2021). Friends or roes-microbial interactions in nature. Biology *10*, 496. https://doi.org/10.3390/biology10060496.
47. Ramoneda, J., Jensen, T.B.N., Price, M.N., Casamayor, E.O., and Fierer, N. (2023). Taxonomic and environmental distribution of bacterial amino acid auxotrophies. Nat. Commun. *14*, 7608. https://doi.org/10.1038/s41467-023-43435-4.
48. Starke, S., Harris, D.M.M., Zimmermann, J., Schuchardt, S., Oumari, M., Frank, D., Bang, C., Rosenstiel, P., Schreiber, S., Frey, N., et al. (2023). Amino acid auxotrophies in human gut bacteria are linked to higher microbiome diversity and long-term stability. ISME J. *17*, 2370–2380. https://doi.org/10.1038/s41396-023-01537-3.
49. van der Wielen, N., Moughan, P.J., and Mensink, M. (2017). Amino acid absorption in the large intestine of humans and porcine models. J. Nutr. *147*, 1493–1498. https://doi.org/10.3945/jn.117.248187.
50. Kramer, J., Özkaya, Ö., and Kümmerli, R. (2020). Bacterial siderophores in community and host interactions. Nat. Rev. Microbiol. *18*, 152–163. https://doi.org/10.1038/s41579-019-0284-4.
51. Kost, C., Patil, K.R., Friedman, J., Garcia, S.L., and Ralser, M. (2023). Metabolic exchanges are ubiquitous in natural microbial communities. Nat. Microbiol. *8*, 2244–2252. https://doi.org/10.1038/s41564-023-01511-x.
52. Machiels, K., Joossens, M., Sabino, J., De Preter, V., Arijs, I., Eeckhaut, V., Ballet, V., Claes, K., Van Immerseel, F., Verbeke, K., et al. (2014). A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. Gut *63*, 1275–1283. https://doi.org/10.1136/gutjnl-2013-304833.
53. Hitch, T.C.A., Riedel, T., Oren, A., Overmann, J., Lawley, T.D., and Clavel, T. (2021). Automated analysis of genomic sequences facilitates high-throughput and comprehensive description of bacteria. ISME Commun. *1*, 16. https://doi.org/10.1038/s43705-021-00017-z.
54. Valentino, V., De Filippis, F., Marotta, R., Pasolli, E., and Ercolini, D. (2024). Genomic features and prevalence of Ruminococcus species in humans are associated with age, lifestyle, and disease. Cell Rep. *43*, 115018. https://doi.org/10.1016/j.celrep.2024.115018.
55. Kim, S.-Y., Woo, S.-Y., Kim, H.-L., Chang, Y., Ryu, S., and Kim, H.-N. (2025). A shotgun metagenomic study identified short-chain fatty acid-producing species and their functions in the gut microbiome of adults with depressive symptoms: Large-scale shotgun sequencing data of the gut microbiota using a cross-sectional design. J. Affect. Disord. *376*, 26–35. https://doi.org/10.1016/j.jad.2025.01.149.
56. Zeng, S., Wang, H., Mu, D., and Wang, S. (2025). The African gut microbiome: A window into the hidden human microbial diversity. Cell Host Microbe *33*, 320–322. https://doi.org/10.1016/j.chom.2025.02.015.
57. Silby, M.W., Winstanley, C., Godfrey, S.A.C., Levy, S.B., and Jackson, R. W. (2011). Pseudomonas genomes: diverse and adaptable. FEMS Microbiol. Rev. *35*, 652–680. https://doi.org/10.1111/j.1574-6976.2011.00269.x.
58. Jiang, D., Goswami, R., Dennis, M., Heimsath, H., Kozlowski, P.A., Ardeshir, A., Van Rompay, K.K.A., De Paris, K., Permar, S.R., and Surana, N.K. (2023). Sutterella and its metabolic pathways positively correlate with vaccine-elicited antibody responses in infant rhesus macaques. Front. Immunol. *14*, 1283343. https://doi.org/10.3389/fimmu.2023.1283343.
59. Ryan, M.P., and Adley, C.C. (2014). Ralstonia spp.: emerging global opportunistic pathogens. Eur. J. Clin. Microbiol. Infect. Dis. *33*, 291–304. https://doi.org/10.1007/s10096-013-1975-9.
60. Tecer, D., Gogus, F., Kalkanci, A., Erdogan, M., Hasanreisoglu, M., Ergin, Ç., Karakan, T., Kozan, R., Coban, S., and Diker, K.S. (2020). Succinivibrionaceae is dominant family in fecal microbiota of Behcet's Syndrome patients with uveitis. PLoS One *15*, e0241691. https://doi.org/10.1371/journal.pone.0241691.
61. Ryan, M.P., Sevjahova, L., Gorman, R., and White, S. (2022). The emergence of the genus Comamonas as important opportunistic pathogens. Pathogens *11*, 1032. https://doi.org/10.3390/pathogens11091032.
62. Talapko, J., Juzbašić, M., Meštrović, T., Matijević, T., Mesarić, D., Katalinić, D., Erić, S., Milostić-Srb, A., Flam, J., and Škrlec, I. (2024). *Aggregatibacter actinomycetemcomitans*: from the oral cavity to the heart valves. Microorganisms *12*, 1451. https://doi.org/10.3390/microorganisms12071451.
63. Ren, X., and Palmer, L.D. (2023). Acinetobacter metabolism in infection and antimicrobial resistance. Infect. Immun. *91*, e0043322. https://doi.org/10.1128/iai.00433-22.
64. Denamur, E., Clermont, O., Bonacorsi, S., and Gordon, D. (2021). The population genetics of pathogenic Escherichia coli. Nat. Rev. Microbiol. *19*, 37–54. https://doi.org/10.1038/s41579-020-0416-x.
65. Hu, J., Amor, D.R., Barbier, M., Bunin, G., and Gore, J. (2022). Emergent phases of ecological diversity and dynamics mapped in microcosms. Science *378*, 85–89. https://doi.org/10.1126/science.abm7841.
66. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.
67. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
68. Li, D., Liu, C.M., Luo, R., Sadakane, K., and Lam, T.W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics *31*, 1674–1676. https://doi.org/10.1093/bioinformatics/btv033.
69. Uritskiy, G.V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome *6*, 158. https://doi.org/10.1186/s40168-018-0541-1.
70. Chklovski, A., Parks, D.H., Woodcroft, B.J., and Tyson, G.W. (2023). CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nat. Methods *20*, 1203–1212. https://doi.org/10.1038/s41592-023-01940-w.
71. Olm, M.R., Brown, C.T., Brooks, B., and Banfield, J.F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J. *11*, 2864–2868. https://doi.org/10.1038/ismej.2017.126.
72. Chaumeil, P.A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics *36*, 1925–1927. https://doi.org/10.1093/bioinformatics/btz848.
73. Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf. *11*, 119. https://doi.org/10.1186/1471-2105-11-119.
74. Arango-Argoty, G., Garner, E., Pruden, A., Heath, L.S., Vikesland, P., and Zhang, L. (2018). DeepARG: a deep learning approach for predicting

antibiotic resistance genes from metagenomic data. Microbiome *6*, 23. https://doi.org/10.1186/s40168-018-0401-z.

75. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. Nat. Methods *12*, 59–60. https://doi.org/10.1038/nmeth.3176.
76. Price, M.N., and Arkin, A.P. (2017). PaperBLAST: Text Mining Papers for Information about Homologs. mSystems *2*, e00039-17, 00039-00017. https://doi.org/10.1128/mSystems.00039-17.
77. Koblitz, J., Reimer, L.C., Pukall, R., and Overmann, J. (2025). Predicting bacterial phenotypic traits through improved machine learning using high-quality, curated datasets. Commun. Biol. *8*, 897. https://doi.org/10.1038/s42003-025-08313-3.
78. Aroney, S.T.N., Newell, R.J.P., Nissen, J.N., Camargo, A.P., Tyson, G.W., and Woodcroft, B.J. (2025). CoverM: read alignment statistics for metagenomics. Bioinformatics *41*, btfa147. https://doi.org/10.1093/bioinformatics/btaf147.
79. Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. PLoS One *11*, e0163962. https://doi.org/10.1371/journal.pone.0163962.
80. Machado, D., Andrejev, S., Tramontano, M., and Patil, K.R. (2018). Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. Nucleic Acids Res. *46*, 7542–7553. https://doi.org/10.1093/nar/gky537.
81. Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336. https://doi.org/10.1038/nmeth.f.303.
82. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
83. RCoreTeam (2013). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://cran.r-project.org/.
84. Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633.
85. Schliep, K.P. (2011). phangorn: phylogenetic analysis in R. Bioinformatics *27*, 592–593. https://doi.org/10.1093/bioinformatics/btq706.
86. Hadfield, J.D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J. Stat. Software *33*, 1–22. https://doi.org/10.18637/jss.v033.i02.
87. Dixon, P. (2003). VEGAN, a package of R functions for community ecology. J. Veg. Sci. *14*, 927–930. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x.
88. Harrell, J.F.E. (2019). Package 'hmisc'. CRAN2018 *2019*, 235–236. https://cran.r-project.org/web/packages/Hmisc/index.html.
89. RCoreTeam (2013). The R Stats Package. R Package Version 4.0.3. https://search.r-project.org/R/refmans/stats/html/stats-package.html.
90. Cutler, D.R., Edwards, T.C., Jr., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., and Lawler, J.J. (2007). Random forests for classification in ecology. Ecology *88*, 2783–2792. https://doi.org/10.1890/07-0539.1.
91. Wickham, H. (2011). ggplot2. WIREs Computational Stats. *3*, 180–185. https://doi.org/10.1002/wics.147.
92. Yin, X., Jiang, X.T., Chai, B., Li, L., Yang, Y., Cole, J.R., Tiedje, J.M., and Zhang, T. (2018). ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. Bioinformatics *34*, 2263–2270. https://doi.org/10.1093/bioinformatics/bty053.
93. Liu, B., Zheng, D., Jin, Q., Chen, L., and Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. *47*, D687–D692. https://doi.org/10.1093/nar/gky1080.
94. Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., Prill, R.J., Tripathi, A., Gibbons, S.M., Ackermann, G., et al. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. Nature *551*, 457–463. https://doi.org/10.1038/nature24621.
95. Liu, Y., Huang, G., Wei, F., and Hu, Y. (2025). Non-negligible role of gut morphology in shaping mammalian gut microbiomes. Sci. China Life Sci. *68*, 2408–2419. https://doi.org/10.1007/s11427-024-2933-1.
96. Amir, A., McDonald, D., Navas-Molina, J.A., Kopylova, E., Morton, J.T., Zech Xu, Z., Kightley, E.P., Thompson, L.R., Hyde, E.R., Gonzalez, A., and Knight, R. (2017). Deblur rapidly resolves single-nucleotide community sequence patterns. mSystems *2*, e00191-16–e00116. https://doi.org/10.1128/mSystems.00191-16.
97. Zheng, K., Liang, D., Wang, X., Han, Y., Griesser, M., Liu, Y., and Fan, P. (2022). Contrasting coloured ventral wings are a visual collision avoidance signal in birds. Proc. Biol. Sci. *289*, 20220678. https://doi.org/10.1098/rspb.2022.0678.
98. Symonds, M.R.E., and Blomberg, S.P. (2014). A Primer on Phylogenetic Generalised Least Squares (Springer). https://doi.org/10.1007/978-3-662-43550-2_5.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
ALFA-SEQ Stool DNA Kit	Magigene	Cat#ADX1050
ALFA-SEQ DNA Library Prep Kit	Magigene	Cat#DL1010
NovaSeq 6000 instrument	Illumina	N/A
Deposited data		
Metagenomes	This study	NCBI: PRJNA1158499
Metabolic models for the 3,867 SGBs	This study	Zenodo: https://doi.org/10.5281/zenodo.17140653
Scripts and relevant data for the main and supplementary figures	This study	Zenodo: https://doi.org/10.5281/zenodo.17139167
Software and algorithms		
fastp v0.23.4	Chen et al.[66]	https://github.com/OpenGene/fastp
bwa v0.7.17	Li and Durbin[67]	https://github.com/lh3/bwa
Megahit v1.2.9	Li et al.[68]	https://github.com/voutcn/megahit
metaWRAP v1.3.2	Uritskiy et al.[69]	https://github.com/bxlab/metaWRAP
CheckM2 v1.0.1	Chklovski et al.[70]	https://github.com/chklovski/CheckM2
dRep v3.5.0	Olm et al.[71]	https://github.com/MrOlm/drep
CompareM v0.1.2	N/A	https://github.com/dparks1134/CompareM
GTDB-Tk v2.4.0	Chaumeil et al.[72]	https://github.com/Ecogenomics/GTDBTk
Prodigal v2.6.3	Hyatt et al.[73]	https://github.com/hyattpd/Prodigal
DeepARG v1.0.2	Arango-Argoty et al.[74]	https://github.com/gaarangoa/deeparg
diamond v2.1.10	Buchfink et al.[75]	https://github.com/bbuchfink/diamond/
PaperBLAST v1.2	Price and Arkin[76]	https://github.com/morgannprice/PaperBLAST
Bacdive-AI v1.0	Koblitz et al.[77]	https://github.com/LeibnizDSMZ/bacdive-AI
CoverM v0.7.0	Aroney et al.[78]	https://github.com/wwood/CoverM
seqkit v2.7.0	Shen et al.[79]	https://github.com/shenwei356/seqkit
CarveMe v1.6.0	Machado et al.[80]	https://github.com/cdanielmachado/carveme
SMETANA v1.0	Zelezniak et al.[21]	https://github.com/cdanielmachado/smetana
QIIME2 v2024.2.0	Caporaso et al.[81]	https://qiime2.org/
BLAST v2.15.0	Altschul et al.[82]	https://www.ncbi.nlm.nih.gov/
R v4.2.3	RCoreTeam[83]	https://cran.r-project.org/
ape v5.6-2	Paradis and Schliep[84]	https://cran.r-project.org/web/packages/ape/index.html
phangorn v2.11.1	Schliep[85]	https://cran.r-project.org/web/packages/phangorn/index.html
MCMCglmm v2.35	Hadfield[86]	https://cran.r-project.org/web/packages/MCMCglmm/index.html
vegan v2.6-4	Dixon[87]	https://cran.r-project.org/web/packages/vegan/index.html
Hmisc v4.2-0	Harrell[88]	https://cran.r-project.org/web/packages/Hmisc/index.html
stats v4.2.3	RCoreTeam[89]	https://search.r-project.org/R/refmans/stats/html/stats-package.html
randomForest v4.7–1.1	Cutler et al.[90]	https://cran.r-project.org/web/packages/randomForest/index.html

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ggplot2 v3.5.1	Wickham[91]	https://cran.r-project.org/web/packages/ggplot2/index.html
Other		
SARG v3.2.1-S	Yin et al.[92]	https://smile.hku.hk/ARGs/Indexing/download
VFDB	Liu et al.[93]	https://www.mgc.ac.cn/VFs/
Amplicon data of NHPs	Amato et al.[36]	ERP016286
EMP data	Thompson et al.[94]	https://earthmicrobiome.org/data-and-code/
GTDB database release 220	Chaumeil et al.[72]	https://data.gtdb.ecogenomic.org/releases/latest/
Primate phylogenetic trees	N/A	https://vertlife.org/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Ethical approval

We carried the field study and collected feces of the sympatric western black crested gibbons and Indochinese gray langurs distributed around the Dazhaizi Gibbon Research Station (100°42 'E, 24°21 'N) in Mt. Wuliang National Nature Reserve, Yunnan, China. In the present study, we focused on the impact of diet, habitat, gut morphology, and host phylogeny on gut microbial metabolic interactions, as the age and sex of other NHPs species were mostly unavailable in previous studies. The age and sex of the two sympatric primate species are thus not relevant to our analyses. During follows, all fecal samples were non-invasively collected without disturbing the gibbons and langurs. All data and samples were obtained legally and undertaken with permissions of Yunnan Wuliangshan Ailaoshan National Nature Reserve, Jingdong Bureau (permit date: 20201205).

METHOD DETAILS

Public metagenome collection

We searched publications in the Scopus database to retrieve articles focusing on the gut microbial metagenomes of non-human primates. We manually excluded studies utilizing primates as experimental animal models. Finally, we obtained a total of 23 published papers that provided 725 metagenomes ≥1 Gb for public access (Table S1). These metagenomes were sequenced from the fecal samples of 51 primate species. We also collected species traits, including the countries where the fecal samples were collected, the dietary types of each NHP species, the living environments, and the gut morphology from the papers. As the dietary data and gut morphology of some species were not provided in the corresponding paper, we further collected these data from IUCN or a recently published paper.[95]

Wild data collection

We collected fecal samples for the sympatric western black crested gibbons and Indochinese gray langurs distributed around the Dazhaizi Gibbon Research Station (100°42 'E, 24°21 'N) in Mt. Wuliang National Nature Reserve, Yunnan, China. The gibbons and langurs in this area have been continuously monitored by our group for nearly twenty years. Three groups of gibbons and a group of langurs have been habituated to allow close observations and fecal sample collection. From April 2021 to January 2022, we carried out our longitudinal field study for the habituated gibbon and langur groups. Once a defection occurred, we collected the feces free of soil and litter contamination into 15 mL sterile tubes with 95% ethanol within five minutes, and then transported to the laboratory where they were stored at −80°C. We finally collected 116 fecal samples for western black crested gibbons (n = 68) and Indochinese gray langurs (n = 48) (Table S1). The total microbial DNA was extracted from the fecal samples using the ALFA-SEQ Stool DNA Kit (Magigene, Guangdong, China) following the manufacturer's instructions. Sequencing libraries were prepared with ALFA-SEQ DNA Library Prep Kit (Magigene, Guangdong, China) and sequenced from 150 bp paired ends on an Illumina NovaSeq 6000 platform (30 Gb for each library).

Genome catalog construction

The 116 sequenced metagenomes, combined with the 725 publicly downloaded metagenomes from the 23 published studies were quality-trimmed, removed with host DNA contamination, and assembled. Specifically, the metagenomic data were quality filtered with fastp v0.23.4 with parameters '-D -n 5 –M 20 –L 75 -Q'.[66] Primate genomes for each of the 53 species or their closest relatives were downloaded from the NCBI website (Table S1). The high-quality metagenomic reads were mapped against the corresponding primate genomes using bwa v0.7.17 to remove contaminant sequences from the hosts.[67] The clean reads for each metagenome were assembled independently using Megahit v1.2.9.[68] The assembled contigs ≥2,500 bp from each metagenome were binned to microbial genomes using metaWRAP v1.3.2 with 'binning -l 2500 -metabat2 -maxbin2 -concoct' and 'bin_refinement -c 50 -x

10' options.[69] Meanwhile, microbial genomes assembled from a previous study regarding the NHPs gut metagenomes were also downloaded (Figure S1A).[37]

The completeness and contamination of all recovered and downloaded microbial genomes were determined using CheckM2 v1.0.1.[70] Genomes with estimated completeness >90% and contamination <5% were retained. The resulted 16,340 high-quality genomes were aggregated and clustered into 3,867 SGBs using dRep v3.5.0 with options 'dereplicate -comp 90 -con 5 -sa 0.95 -cm larger -nc 0.3 –p 40 -comW 1 -conW 4 -strW 0 -N50W 0 -sizeW 0 -centW 0'.[71] To estimate genomic relatedness between SGBs, AAI was calculated using the CompareM (v0.1.2) AAI workflow (https://github.com/dparks1134/CompareM).

Taxonomic and functional annotation

For the 3,867 SGBs, taxonomic assignment was performed with GTDB-Tk v2.4.0 (database release R09-RS220).[72] Open-reading-frames (ORFs) the SGBs were predicted using Prodigal v2.6.3 with the parameters set as '-p meta -g 11 -f gff -q -m'.[73] ARGs were annotated using two methods as follows. First, DeepARG v1.0.2 was applied on the predicted protein sequences using the 'LS' model with options '–min-prob 0.8 –arg-alignment-identity 50 –arg-alignment-evalue 1e-10 –arg-alignment-overlap 0.8'.[74] Second, sequence alignment was conducted using diamond v2.1.10 'blastp' command to align predicted protein sequences against the SARG (v3.2.1-S, accessed on 2025.01.11) with options '–evalue 1e-10 –query-cover 80 –subject-cover 80 –id 50'.[75,92] VFs were annotated by aligning the predicted protein sequences to the reference sequence in the core set of Virulence Factors of Pathogenic Bacteria Database (VFDB) using diamond v2.1.10 'blastp' command with options '–evalue 1e-10 –query-cover 80 –subject-cover 80 –id 50'.[75,93] To explore the capability of amino acids biosynthesis and common carbon sources utilization for genomes, the GapMind code available in PaperBLAST v1.2 was used,[76] and pathways that had all high-confidence steps were retained. To predict the traits of oxygen requirements for genomes, bacdive-AI v1.0 was used and the confidence values for anaerobic and aerobic traits were obtained.[77]

Computation of MRO and MIP scores

Since the 841 metagenomic datasets varied in reads length and number, we firstly mapped the 3,867 redundant SGBs to the raw metagenomes using CoverM v0.7.0 using the parameters '-p minimap2-sr –min-read-percent-identity 0.95 –min-read-aligned-percent 0.75 -m trimmed_mean –trim-min 0.05 –trim-max 0.95 –min-covered-fraction 0' to retrieve the gut microbial community composition.[78] We then selected the dominant (relative abundance >0.1%) SGBs in each sample to estimate the community-level MRO and MIP scores. To assess the impact of the sequencing depth on the recovery of dominant SGBs, raw metagenomes with largest size in each of the nine primate families were selected and randomly subsampled with different proportions (0.8, 0.6, 0.4, and 0.2) using seqkit v2.7.0 to simulate metagenomes with various sequencing depth.[79] The 3,867 SGBs were then mapped to the subsampled metagenomes to assess the recall rate of dominant SGBs detected in the raw metagenomes. Finally, to maximize the recovery of dominant SGBs and minimize sampling biases, 184 raw metagenomes with microbial richness $\geq$10 and reads mappability $\geq$50% were selected to calculate community MRO/MIP.

Genome-scale metabolic model for each of the 3,867 SGBs was reconstructed from its protein fasta file using CarveMe v1.6.0.[80] The gap-filling process was performed with nutrient-rich and anaerobic conditions (i.e., -g LB[-O2],M9[-O2]), given the eutrophic and low-oxygen conditions in the gut. Next, the MRO and MIP scores of dominant SGBs in a given sample was assessed using the global mode of SMETANA v1.0.[21] Meanwhile, to estimate the competition/cooperation strength for each SGB, the MRO/MIP scores between a single SGB and all other SGBs were calculate and averaged.[21] The competition/cooperation strength for a specific phylum was thus calculated as the mean competition/cooperation strength of SGBs assigned to that phylum. To elucidate the input and output compounds of SGBs, contructed metabolic models were converted into directed graphs using the the R package xml2 v1.3.6.[83] The reactants and products for exchange and sink reactions were subsequently identified from these graphs. Finally, the input and output compounds for the SGBs were defined as the identified reactants and products, excluding inorganic compounds (e.g., water, nitrogen, and oxygen) which are assumed to be consistently present in the external gut environment.[21]

We generated random communities for the phyla that exhibited highest number of SGBs in our datasets (Figure S4D). In detail, for a random community of size n, we initially randomly selected SGBs within a given phyla with numbers (size = k) not exceeding n. Subsequently, we randomly selected additional SGBs (size = n - k) from the remaining SGBs assigned to other phyla. The MRO and MIP scores of the simulated communities then were calculated using the global mode of SMETANA v1.0.[80]

Amplicon data recruitment and processing

The EMP data, including the community composition, representative sequences, and sample types were recruited from a previous study.[94] To investigate the potential associations between gut microbial metabolic interactions and host health, we downloaded amplicon dataset (ERP016286, https://qiita.ucsd.edu/study/description/1453) from a previously published study investigating the gut microbiota differences between healthy and unhealthy NHPs.[36] All the downloaded 16S rRNA gene sequences were processed using Quantitative Insights into Microbial Ecology (QIIME2 v 2024.2.0) defaults and classified into ASVs using the deblur workflow.[81,96] Samples were rarefied at 10,000 sequences per sample to remove those with lower sequencing depth and normalize read counts across samples.

To assign the GTDB taxonomy or genomes to the representative 16S rRNA gene sequences from the EMP project or the NHPs amplicon data, all 113,104 representative genomes were downloaded from GTDB database (release 220). The representative 16S

rRNA gene sequences were mapped to these genomes using BLAST v2.15.0 (BLASTn) with a 99% identify threshold and 95% alignment coverage.[82] For the NHPs amplicon data, 16S rRNA gene sequences were mapped to both the GTDB genomes and NPG-MG genomes. If multiple genomes were found for a ASV sequence, the genome with highest alignment identify and length was selected.[20]

MCMCglmm analysis

We considered species traits, including diet, living environment, gut morphology, location, and phylogeny that might affect the primate gut microbial metabolic interactions (Tables S1 and S3). We downloaded 100 node-dated phylogenetic trees of the 53 species from VertLife.org. For the species that were not included in the VertLife database, closest relatives of them were used as alternatives (Table S1). We then transformed the 100 non-ultrametric trees into ultrametric trees using the '*chronos'* function in ape package v5.6-2 in R v4.2.3.[83,84] We also computed the maximum clade credibility tree from the 100 ultrametric trees for graphical view using the '*maxCladeCred*' function in R package phangorn v2.11.1.[83,85]

To control the effect of phylogeny, we conducted phylogenetic comparative analyses using the R package MCMCglmm v2.35 with the fixed effects included the diet, habitat, and distance from the equator.[86] Both species and country were included as random effects. We standardized all numeric variables with the 'decostand' function in R package vegan v2.6-4.[87] We used a Gaussian distribution with a parameter expanded prior R = list(V = 1, nu = 0.002), G = list(G1 = G2 = list[V = 1, nu = 1, alpha.mu = 0, alpha.V = 1000]).[97] We ran all models with the 100 ultrametric trees for 160,000 iterations and with a burn-in of 10,000 iterations and a thinning interval of 30. Convergence was assessed with Gelman-Rubin diagnostics, confirming high estimated sample sizes, and visualization of scale reduction factor. We then estimated the posterior probability of the phylogenetic signal for MCMCglmm models by dividing the phylogenetic variance component by the variance components.[98]

Statistical analyses

All statistical analyses were performed with various packages within the statistical program R v4.2.3.[83] These analyses include correlation estimation (Hmisc v4.2-0),[88] significance comparison and correction (stats v4.2.3),[89] random forest model construction (randomForest v4.7–1.1),[90] and data visualization (ggplot2 v3.5.1).[91] Relationships between different factors across samples were investigated by Pearson's correlation analyses using '*rcorr*' function in Hmisc v4.2-0.[88] Comparison of the values between different groups was performed using '*pairwise.wilcox.test*' function (multiple-group comparison) or '*wilcox.test*' function (two-group comparison) (stats v4.2.3).[89] The p values were adjusted by multiple testing corrections using the Benjamini and Hochberg false discovery rate (FDR) controlling procedure (stats v4.2.3).[89] An FDR-corrected $p < 0.05$ was considered statistically significant for comparing. The random forest with leave-one-out cross-validation was performed using '*randomForest*' function based on the presence-absence matrix of key microbiota to conduct classification analysis.[90] All results were visualized with boxplots, line plots, bar plots, or heatmaps plotted with the '*ggplot2*' function in ggplot2 v3.5.1.[91]

QUANTIFICATION AND STATISTICAL ANALYSIS

For the metagenomic analyses, the gut microbial richness and abundance were calculated by mapping the 3,867 redundant SGBs to the raw metagenomes using CoverM v0.7.0 (Figure 1).[78] The dominant (relative abundance >0.1%) SGBs in each sample were used to assess the MRO and MIP scores using the global mode of SMETANA v1.0 (Figure 2).[21] MCMCglmm analysis was performed to examine the impact of different factors on the gut microbial metabolic interactions (Figure 2). For the statistical analyses (Figures 2, 3, 4, and 5), unless otherwise stated, comparison of the values between different groups was performed using '*pairwise.wilcox.test*' function (multiple-group comparison) or '*wilcox.test*' function (two-group comparison) (stats v4.2.3).[89] Other details on statistical analysis can be found in Figure legends.