Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

First Metabarcoding Characterization of Seabird Guano Microbiomes Along the Peruvian Coast: Environmental Filtering Shapes Taxonomically Simplified but Functionally Stable Communities

Roberto Cosme

rcosme@inia.gob.pe

Instituto Nacional de Innovación Agraria

Héctor Cántaro-Segura

Universidad Nacional Agraria La Molina

Enrique Meseth

Instituto Nacional de Innovación Agraria

Anderson Calixto

Instituto Nacional de Innovación Agraria

María del Pilar López

Instituto Nacional de Innovación Agraria

Vicente Ruiz

Instituto Nacional de Innovación Agraria

Luz Palomino

Instituto Nacional de Innovación Agraria

Arturo Calderón

Programa de Desarrollo Productivo Agrario Rural

Licia Salvatierra

Instituto Tecnológico de la Producción

Renzo Valdez

Universidad Nacional de Barranca

Susan Medina

Scientific University of the South

Article

Keywords: seabird guano, metabarcoding, 16S rRNA, ITS, microbial diversity, environmental filtering, biofertilization, Leucocarbo bougainvillii, Sula variegata, Peruvian coastline

Posted Date: May 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9316068/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Seabird guano from the islands off the Peruvian coast constitutes an organic fertilizer resource of recognized strategic importance; however, the microbial communities associated with this substrate had remained entirely uncharacterized. The present study provides the first comprehensive, culture-independent characterization of bacterial and fungal communities in guano collected from 13 guano-producing zones along the Peruvian coastline, encompassing deposits from the Guanay cormorant (*Leucocarbo bougainvillii*) and the Peruvian booby (*Sula variegata*), the two dominant and representative guano-producing bird species. Using 16S rRNA and ITS amplicon metabarcoding, combined with physicochemical analysis and functional prediction through FAPROTAX, and FUNGuild, we found that bacterial communities were consistently dominated by Bacillota (> 90%), particularly the order Bacillales, with halotolerant and endospore-forming genera, among which *Saccharomonospora* was notably abundant; this genus exhibited the strongest positive correlation with potassium content (+ 0.80, $p < 0.05$) and a significant negative correlation with organic carbon (− 0.79, $p < 0.05$), positioning it as a potential bioindicator of the chemical maturation of the substrate. Fungal communities were likewise dominated by Ascomycota (> 95%), with *Aspergillus* as the only genus showing a significant positive differential-abundance signal in Guanay cormorant guano relative to Peruvian booby guano. Within this exploratory framework, although the overall composition of bacterial and fungal communities did not differ significantly between the two bird species, differential abundance analysis identified specific bacterial taxa, including *Pseudomonas*, *Atopostipes*, and *Lentibacillus*, as significantly enriched in Guanay cormorant guano, revealing microbial differences that, while subtle, are ecologically relevant across host species. The predicted functional profile suggested a broadly stable metabolic repertoire oriented towards organic matter mineralization and nitrogen transformation across samples. These findings provide a first microbiological baseline for the evidence-based management of Peruvian island guano as a sustainable biofertilization resource.

Introduction

Seabird guano represents one of the most nutrient-rich organic fertilizers known, and its large-scale exploitation during the nineteenth century fundamentally transformed global agricultural production [1,2]. Along the Peruvian coast, the convergence of the cold, highly productive upwelling system of the Humboldt Current with the nesting behavior of two emblematic guano-producing seabirds, the Guanay cormorant (*Leucocarbo bougainvillii*) and the Peruvian booby (*Sula variegata*), gives rise to extraordinary accumulations of nitrogen-, phosphorus-, and potassium-rich excreta on rocky islands and coastal headlands [1,3,4]. These deposits, managed as protected natural reserves by the Peruvian state, constitute a strategic biological resource of both ecological and agricultural significance [5,6]. The renewed global interest in natural organic fertilizers driven by growing evidence of the environmental costs of synthetic inputs and recent disruptions to fertilizer supply chains has repositioned Peruvian island guano as a key asset for sustainable food production [7,8]. Field experiments conducted in Peru have confirmed that

island guano application substantially improves soil fertility and crop yield in degraded quinoa (*Chenopodium quinoa*) and cañihua (*Chenopodium pallidicaule*) cropping systems [6,9–12].

Beyond its well-established role as a macronutrient reservoir, seabird guano constitutes a highly dynamic microbial habitat. The sustained deposition of excreta rich in uric acid, proteins, and lipids creates chemically extreme microenvironments characterized by high organic matter content, elevated nitrogen concentrations, variable salinity, and fluctuating pH [13,14]. Under these conditions, microbial communities are not passive inhabitants but active agents driving organic matter decomposition, nitrogen and phosphorus mineralization, and carbon compound transformation through biogeochemical pathways of direct relevance to soil fertility [15–17]. The chemical composition of guano therefore reflects a tight interplay between host bird physiology, dietary inputs, and the metabolic activity of the resident microbiome, with microbial assemblages being shaped by both host-derived selective pressures and environmental factors [14,18,19].

Microbial diversity in guano deposits has been characterized primarily in cave-dwelling bat species and polar seabird colonies, where high-throughput sequencing has revealed phylogenetically diverse assemblages dominated by Proteobacteria, Firmicutes, Actinobacteria, and Bacteroidetes [15,16,18,20]. In Antarctic systems, guano deposition by the Adélie penguin (*Pygoscelis adeliae*) has been shown to generate ornithogenic soils characterized by elevated ammonia and uric acid derivatives that sustain microbial communities adapted to chemically harsh conditions[13]. More recently, evidence from temperate coastal environments has demonstrated that seabird guano inputs significantly alter microbial community composition and phytoplankton–bacteria interactions in adjacent marine systems [21]. Despite this growing body of knowledge, the microbial communities associated with guano from Peruvian island seabirds have not been characterized using culture-independent, high-resolution approaches. This gap is particularly significant given the ecologically distinctive character of these island ecosystems, their multi-century history of intense avian occupation, and the demonstrated value of their guano as an organic fertilizer resource.

The advent of culture-independent, high-throughput sequencing technologies has fundamentally transformed the study of environmental microbial communities, enabling comprehensive characterization of both abundant and rare taxa in complex substrates [22,23]. Amplicon metabarcoding using hypervariable regions of the 16S rRNA gene and the internal transcribed spacer (ITS) region provides fine and parallel taxonomic resolution of bacterial and fungal diversity [24,25]. These complementary markers offer distinct ecological perspectives: bacterial communities tend to exhibit functional redundancy in nutrient-rich substrates, whereas fungal assemblages are generally more sensitive to microenvironmental variability and resource heterogeneity, making their combined application particularly informative in chemically complex organic substrates such as guano [26–28]. Physicochemical gradients — particularly pH and salinity — rank among the most powerful determinants of microbial community structure and diversity at both local and regional scales [29,30]. In saline or hypersaline environments, osmotic stress acts as a potent environmental filter, selecting for taxa with

physiological adaptations including compatible solute synthesis, ion exclusion mechanisms, and spore formation [31]. In guano-derived substrates, where electrical conductivity and pH can vary considerably across deposits from different islands, bird species, and accumulation stages, abiotic gradients are expected to exert a dominant structuring influence on microbial communities, potentially overriding the effect of host identity. Understanding these environmental drivers is therefore fundamental to interpreting spatial variation in community composition and to predicting the agronomic implications of applying guano from diverse sources to agricultural soils.

The integration of microbiome characterization into the evaluation of organic fertilizers is increasingly recognized as a prerequisite for their evidence-based agronomic use [8,16]. Guano harboring microbial consortia capable of biological nitrogen fixation, phosphate solubilization, and plant growth promotion would provide dual agronomic benefits: as a direct nutrient source and as a vehicle for inoculating beneficial microorganisms into the soil [7,32]. Conversely, guano batches dominated by opportunistic pathogens or soil-degrading taxa could pose risks if applied without prior microbiological assessment [33]. Metabarcoding provides a scalable, rapid, and reproducible framework for this type of quality assessment, enabling the selection of guano batches with optimal microbial profiles for sustainable biofertilization strategies in Peruvian agroecosystems [9,16].

In this context, the present study provides the first comprehensive, culture-independent characterization of bacterial and fungal communities in guano collected from 13 guano-producing zones along the Peruvian coastline, encompassing deposits from the two dominant bird species (*L. bougainvillii* and *S. variegata*). Using 16S rRNA and ITS amplicon metabarcoding, we described alpha and beta diversity patterns, quantified shared and exclusive microbial taxa between bird-associated guano types, and explored the relationship between physicochemical parameters and community structure through redundancy analysis. Additionally, functional profiles were inferred to characterize the predicted metabolic potential of the microbiome. Given the composite-sample design, the study was intended to generate a descriptive baseline and exploratory hypotheses regarding the relative importance of abiotic gradients and host identity.

Results

Relative abundance of bacterial and fungal communities

Taxonomic composition analysis of the bacterial microbiome associated with island guano revealed a consistently dominant structure comprising a small number of highly adapted lineages, with concordant patterns across the different taxonomic levels evaluated (Figs. 1 and S2).

At the phylum level, Bacillota was the absolutely dominant taxon across all samples analyzed, with relative abundances exceeding 90% in the majority of the guano localities evaluated (Figure S2). This pattern was remarkably homogeneous across sites and bird species, with marginal variation attributable to the minor presence of other phyla. Secondary phyla identified included Actinomycetota,

Pseudomonadota, Halobacteriota, and Verrucomicrobiota, whose combined representation did not exceed 10% in any sample. The pronounced dominance of Bacillota throughout the entire Peruvian coastline suggests the existence of strong environmental filtering that consistently favors this lineage in the island guano ecosystem.

At the order level, Bacillales was the predominant taxon across all samples, representing between 25% and 95% of relative abundance depending on the locality (Fig. 1a). Peptostreptococcales–Tissierellales constituted the second most abundant order at the majority of sites, with notable variation among localities. Additional orders detected with relevant presence in at least some samples included Clostridiales, Lactobacillales, Propionibacteriales, and Pseudomonadales, although generally at representations below 10%. The samples from Guañape Norte (Booby) and Guañape Sur (Booby) exhibited a higher relative proportion of Peptostreptococcales–Tissierellales compared to other localities, suggesting local variation in bacterial community structure.

At the family level, Bacillaceae was the dominant family in virtually all samples, followed by Peptostreptococcaceae and Gottschalkiaceae as consistent secondary components (Fig. 1b). Marinococcaceae showed notable presence in some localities, particularly in Guanay cormorant samples. The presence of Halomonadaceae, although minor, is ecologically relevant given its well-known tolerance to high-salinity conditions. A proportion of ASVs was not assigned at the family level and was classified as Unclassified, which may reflect the presence of bacterial lineages that are poorly represented in current reference databases.

At the genus level, the most representative taxa were *Paeniclostridium*, *Virgibacillus*, *Bacillus*, *Lentibacillus*, and *Neobacillus*, with variation in their relative abundance across localities and bird species (Fig. 1c). *Paeniclostridium* showed consistent and prominent presence across multiple sites, while *Virgibacillus* and *Lentibacillus* — genera recognized for their halotolerance — exhibited variable but ecologically relevant abundances. *Bacillus* and *Neobacillus*, both characterized by their sporulation capacity and resistance to environmental stress, were likewise prevalent across the majority of samples. Anaerobic fermentative genera such as *Tissierella*, *Gottschalkia*, and *Atopostipes* were present in smaller but consistent proportions, suggesting the existence of microenvironments with heterogeneous redox conditions within the guano substrate. Taken together, the genus-level composition reflects a community dominated by spore-forming, halotolerant, and metabolically versatile organisms, consistent with the extreme physicochemical conditions characteristic of guano accumulated in exposed insular environments.

Taxonomic composition analysis of the fungal microbiome associated with island guano revealed an even more simplified structure than that observed in bacterial communities, characterized by the near-exclusive dominance of a single phylum and a marked concentration of diversity within few orders and families (Figs. 2 and S3).

At the phylum level, Ascomycota dominated absolutely across all samples analyzed, with relative abundances exceeding 95% in virtually all guano localities evaluated (Figure S3). Minor phyla detected

included Chytridiomycota, Basidiomycota, Mortierellomycota, Mucoromycota, Calcarisporiellomycota, and unassigned sequences classified as Fungi_phy_Incertae_sedis, all with representations below 5% across all samples. The most notable exception was the Isla Macabí (Guanay) sample, where Basidiomycota reached a slightly higher proportion compared to the remaining localities, though without exceeding a marginal level.

At the order level, Eurotiales was the overwhelmingly dominant taxon across all samples, representing between 75% and 98% of relative abundance depending on the locality (Fig. 2a). Secondary orders detected with relevant representation in at least some localities included Mortierellales, Saccharomycetales, Sordariales, Pleosporales, and Onygenales, all generally at abundances below 10%. The Chincha Norte (Guanay) sample stood out for exhibiting a higher relative proportion of secondary orders — particularly Mortierellales and Saccharomycetales — compared to the remaining localities, suggesting local microenvironmental conditions that may favor greater fungal order diversity at that site. The Punta San Juan (Guanay) sample likewise displayed a slightly more heterogeneous order-level composition.

At the family level, Aspergillaceae was the absolutely dominant family in the majority of samples, with relative abundances exceeding 90% of the total fungal community in several localities (Fig. 2b). Mortierellaceae constituted the most frequent secondary family, with notable presence in Guañape Sur (Guanay) and Chincha Norte (Guanay). Other families detected at minor but recurring levels included Chaetomiaceae, Debaryomycetaceae, Sporormiaceae, and Rhizophlyctidaceae. Ascobolaceae showed sporadic presence in some samples, constituting a relevant indicator of substrates rich in organic matter of fecal origin. A variable proportion of fungal ASVs remained unassigned at the family level (Unclassified), particularly in the Chincha Norte (Guanay) and Isla Santa Rosa (Guanay) samples, which may reflect the presence of fungal lineages poorly represented in current reference databases.

At the genus level, the mycobiome structure exhibited its greatest degree of variability across localities, although with a general dominance of genera belonging to Aspergillaceae (Fig. 2c). The unassigned component (Unclassified) represented a substantial fraction in the majority of samples, suggesting considerable cryptic diversity at the genus level that current databases are unable to fully resolve. Among the identified genera, *Mortierella* showed notable relative abundances in Guañape Sur (Guanay) and Isla Macabí (Guanay), reaching proportions between 25% and 50% at these localities. *Alternaria*, *Debaryomyces*, *Chrysosporium*, *Penicillium*, *Aspergillus*, and *Preussia* were likewise detected with variable distributions across sites. The genus *Rhizophlyctis*, belonging to the phylum Chytridiomycota, had sporadic presence in some samples, being ecologically relevant for its role in the degradation of complex polysaccharides. Genera such as *Debaryomyces* and *Chrysosporium* are recognized for their tolerance to high-salinity and desiccation conditions, consistent with the physicochemical characteristics of insular guano. Taken together, the genus-level fungal composition reflects a community dominated by opportunistic saprotrophs with broad osmotic tolerance and the capacity to exploit concentrated organic substrates, with a more pronounced differentiation among localities than that observed in bacterial communities.

The relative taxonomic composition and differential abundance of bacterial communities were compared between Guanay cormorant (*L. bougainvillii*) and Peruvian booby (*S. variegata*) guano at multiple taxonomic levels using the 16S rRNA marker (Fig. 3).

At the phylum level, composition was highly similar between both bird species (Fig. 3a). Bacillota dominated in both groups at proportions approaching 95% of total relative abundance. Pseudomonadota and Actinomycetota were present as minor secondary phyla at comparable proportions between Guanay cormorant and Peruvian booby. At the order level, Bacillales was the dominant order in both species, with a slightly higher proportion in Guanay cormorant guano relative to Peruvian booby (Fig. 3b). Peptostreptococcales–Tissierellales constituted the second most abundant order in both groups, with proportionally greater representation in Peruvian booby. Clostridiales, Lactobacillales, Pseudomonadales, and Halobacterales were present as minor components in both species. At the family level, Bacillaceae was the dominant family in both species, with a slightly higher relative proportion in Guanay cormorant (Fig. 3c). Peptostreptococcaceae showed proportionally greater representation in Peruvian booby. Marinococcaceae, Gottschalkiaceae, and Pseudomonadaceae were present at low and comparable proportions between both groups. At the genus level, variations in relative abundances were observed between Guanay cormorant and Peruvian booby within the same set of dominant taxa (Fig. 3d). In Guanay cormorant guano, relatively higher proportions of *Atopostipes*, *Lentibacillus*, and *Tissierella* were notable, whereas in Peruvian booby guano, *Paeniclostridium*, *Neobacillus*, and *Bacillus* showed comparatively higher proportions. *Virgibacillus*, *Pseudomonas*, and the unclassified component (Unclassified) were present in both groups.

Exploratory differential abundance analysis using DESeq2 highlighted a small set of bacterial taxa with positive log2 fold-change values in Guanay-associated samples relative to Peruvian booby-associated samples (Figs. 3e and 3f). The taxa with the highest positive $\log_2$ fold-change values were unclassified bacteria (Unclassified bacteria), *Pseudomonas*, *Atopostipes*, *Lentibacillus*, *Candidatus* Arthromitus, *Neobacillus*, *Tyzzerella*, and *Anaerillum* (Fig. 3e). The volcano plot confirmed that *Alternaria* and *Pseudomonas* exhibited the highest $\log_2$ fold-change values with statistical significance, while *Lentibacillus* showed significant enrichment with a moderate $\log_2$ fold-change (Fig. 3f). The vast majority of detected taxa showed no significant differences between bird species, represented as grey points in the volcano plot.

The relative taxonomic composition and differential abundance of fungal communities were compared between Guanay cormorant (*L. bougainvillii*) and Peruvian booby (*S. variegata*) guano at multiple taxonomic levels using the ITS marker (Fig. 4).

At the phylum level, Ascomycota dominated in both bird species at proportions exceeding 95% of total relative abundance (Fig. 4a). Mortierellomycota and Chytridiomycota were present as minor secondary phyla in both groups, with slightly higher representations in Guanay cormorant relative to Peruvian booby. Basidiomycota, Calcarisporiellomycota, Mucoromycota, and Fungi_phy_Incertae_sedis were detected at marginal proportions in both species.

At the order level, Eurotiales was the dominant order in both species, with a visibly higher relative proportion in Peruvian booby compared to Guanay cormorant (Fig. 4b). Secondary orders showed greater relative representation in Guanay cormorant than in Peruvian booby, including Mortierellales, Saccharomycetales, Sordariales, Pleosporales, Onygenales, Rhizophlyctidales, Pezizales, and Amphisphaeriales. In Peruvian booby, the unclassified fraction (Unclassified) was greater than in Guanay cormorant at this taxonomic level.

At the family level, Aspergillaceae was the dominant family in both species, with a higher relative proportion in Peruvian booby than in Guanay cormorant (Fig. 4c). Mortierellaceae and Debaryomycetaceae showed relatively higher proportions in Guanay cormorant. Chaetomiaceae, Sporormiaceae, Rhizophlyctidaceae, Pleosporaceae, Onygenaceae, Ascobolaceae, and Spiromastigaceae were present at minor representations, with greater relative presence in Guanay cormorant than in Peruvian booby.

At the genus level, the unclassified component (Unclassified) represented the most abundant fraction in both species (Fig. 4d). *Alternaria* was the identified genus with the highest relative abundance in both groups. *Aspergillus* showed a visibly higher proportion in Peruvian booby compared to Guanay cormorant. *Mortierella*, *Chrysosporium*, *Debaryomyces*, *Penicillium*, *Preussia*, *Ascobolus*, *Rhizophlyctis*, *Spiromastigoides*, and *Acrophialophora* were detected at variable proportions, with greater relative representation in Guanay cormorant than in Peruvian booby. Exploratory differential abundance analysis using DESeq2 identified only a single fungal taxon with a significant between-group signal (Figs. 4e and 4f). *Aspergillus* was the only genus showing a significant difference, with a positive $\log_2$ fold-change of approximately 20, indicating enrichment in Guanay cormorant relative to Peruvian booby (Fig. 4e). The volcano plot confirmed that *Aspergillus* was the sole-colored point with statistical significance, located in the upper right quadrant with a $-\log_{10}$(p_adj) value exceeding 6 (Fig. 4f). All remaining detected fungal taxa showed no significant differences between bird species, represented as grey points concentrated around the central axis of the volcano plot.

Alpha diversity analysis

Alpha diversity of bacterial and fungal communities was assessed using the Observed, Chao1, Shannon, and Simpson indices, comparing results between Guanay cormorant and Peruvian booby guano as well as across guano localities (Fig. 5).

Exploratory comparisons of alpha diversity between Guanay cormorant and Peruvian booby composite samples did not reveal clear between-group patterns in any of the indices evaluated (Fig. 5). The Observed index yielded p = 0.24, Chao1 p = 0.24, Shannon p = 0.5, and Simpson p = 0.4. Across all indices, the medians of both groups were similar, although value dispersion was greater in Guanay cormorant than in Peruvian booby. At the individual sample level, Guañape Norte (Booby) and Guañape Sur (Booby) exhibited the highest Observed and Chao1 diversity values among all localities evaluated, while Isla Macabí (Guanay) and Chincha Norte recorded the lowest (Fig. 5b). For the Shannon index, Isla

Santa (Guanay) and Punta Coles (Guanay) showed the highest values, whereas Guañape Norte (Guanay) recorded the lowest. For the Simpson index, Isla Santa (Guanay) and Punta Coles (Guanay) showed the highest values, while Guañape Norte (Guanay) presented the lowest value among all samples. Comparison of fungal alpha diversity between Guanay cormorant and Peruvian booby likewise revealed no statistically significant differences in any of the indices evaluated (Fig. 5c). The Observed index yielded $p = 0.78$, Chao1 $p = 0.58$, Shannon $p = 0.12$, and Simpson $p = 0.052$, the latter approaching the threshold of statistical significance. Across all indices, value dispersion was greater in Guanay cormorant than in Peruvian booby. At the individual sample level, Isla Pescadores (Guanay) recorded the lowest values for the Observed, Chao1, Shannon, and Simpson indices, making it the sample with the lowest fungal diversity among all localities evaluated (Fig. 5d). Isla Macabí (Guanay) likewise exhibited low values for the Observed and Chao1 indices. Guañape Sur (Guanay) and Isla Mazorca (Guanay) recorded the highest Observed and Chao1 diversity values. For the Shannon index, Chincha Norte (Guanay) presented the highest value, while Punta San Juan (Guanay) recorded the lowest among Guanay cormorant samples. For the Simpson index, Guañape Norte (Booby) and Guañape Sur (Booby) showed values comparable to the Guanay cormorant samples with the highest diversity.

Beta diversity analysis

Beta diversity of bacterial and fungal communities was assessed through multivariate ordination analyses, including NMDS and PCoA with Bray–Curtis and Jaccard dissimilarity metrics, comparing sample distribution according to bird species and geographic region of origin (Fig. 6).

NMDS analysis based on Bray–Curtis dissimilarity for bacterial communities yielded a stress value of 0.096, indicating an acceptable two-dimensional ordination fit (Fig. 6a). In ordination space, Guanay cormorant and Peruvian booby samples showed broad overlap, with no clear separation between groups by bird species. Guañape Norte (Booby) and Guañape Sur (Booby) samples were positioned in regions more distant from the center of ordination space relative to the remaining samples. Isla Santa Rosa (Guanay) and Punta San Juan (Guanay) samples were located in the upper quadrant of ordination space, separated from the northern samples.

Bray–Curtis-based PCoA explained 34.1% of variance on the PCoA1 axis and 15.9% on the PCoA2 axis, accounting for a cumulative 50.0% of total variance across the first two axes (Fig. 6b). Guanay cormorant and Peruvian booby samples showed an overlapping distribution in ordination space, with the exception of Guañape Norte (Booby), which was clearly positioned apart from the remaining samples along the negative PCoA1 axis. The dashed ellipse encompassed the majority of Guanay cormorant samples, reflecting moderate dispersion among localities. Jaccard-based PCoA explained 14.8% of variance on PCoA1 and 12.5% on PCoA2 (Fig. 6c). Sample distribution in this space was similar to that observed with Bray–Curtis, with overlap between Guanay cormorant and Peruvian booby and distant positioning of Guañape Norte (Booby) and Guañape Sur (Booby) samples.

The PCoA colored by geographic region revealed a partial clustering tendency according to sample provenance (Fig. 6d). Northern region samples tended to position toward more negative PCoA1 values, while southern and central region samples distributed toward positive values along that axis. Guañape Norte (Booby) samples were located at the most extreme negative end of the PCoA1 axis, clearly separated from the rest.

NMDS analysis based on Bray–Curtis dissimilarity for fungal communities yielded a stress value of 0.063, indicating an excellent two-dimensional ordination fit (Fig. 6e). Guanay cormorant and Peruvian booby samples showed a dispersed and overlapping distribution in ordination space, with no evident clustering by bird species. The Isla Pescadores (Guanay) sample was clearly positioned apart from the rest in the upper left quadrant, while Isla Santa (Guanay) also showed a peripheral position in ordination space.

Bray–Curtis-based PCoA for fungi explained 48% of variance on PCoA1 and 31.6% on PCoA2, accounting for a cumulative 79.6% of total variance across the first two axes (Fig. 6f). The Isla Pescadores (Guanay) sample occupied an extreme position along the positive PCoA2 axis, clearly separated from all other samples. The remaining samples showed a relatively clustered distribution around the ordination origin, with overlap between Guanay cormorant and Peruvian booby. Jaccard-based PCoA explained 30.4% of variance on PCoA1 and 16.3% on PCoA2 (Fig. 6g). The Isla Pescadores (Guanay) sample was again clearly peripheral, this time along the negative PCoA1 axis. Isla Macabí (Guanay) also showed a position distant from the remaining samples in this ordination space. The other samples displayed a clustered distribution with overlap between Guanay cormorant and Peruvian booby.

The PCoA colored by geographic region for fungi showed a distribution without evident geographic clustering along the PCoA1 axis (Fig. 6h). However, central region samples — particularly Isla Pescadores (Guanay) and Isla Santa (Guanay) — were positioned at higher PCoA2 values, differentiating them from northern and southern region samples, which distributed toward lower values along that axis.

Shared ASVs and taxonomic differentiation

The distribution of shared and exclusive amplicon sequence variants (ASVs) between Guanay cormorant and Peruvian booby guano was assessed using Venn diagrams for the 16S rRNA and ITS markers (Fig. 7).

The total number of bacterial ASVs detected was 2,190 (Fig. 7a). Of these, 1,180 ASVs were exclusive to Guanay cormorant guano, representing 53.9% of the total. Peruvian booby-exclusive ASVs numbered 535, corresponding to 24.4% of the total. ASVs shared between both bird species numbered 475, equivalent to 21.7% of the total bacterial ASVs detected. The Guanay cormorant circle was notably larger than that of Peruvian booby, reflecting the greater number of exclusive ASVs in this group.

The total number of fungal ASVs detected was 318 (Fig. 7b). Of these, 176 ASVs were exclusive to Guanay cormorant guano, representing 55.3% of the total. Peruvian booby-exclusive ASVs numbered 55,

corresponding to 17.3% of the total. ASVs shared between both bird species numbered 87, equivalent to 27.4% of the total fungal ASVs detected. As observed for bacteria, the Guanay cormorant circle was considerably larger than that of Peruvian booby, reflecting greater exclusive ASV richness in Guanay cormorant guano. In proportional terms, the shared fungal core represented a larger fraction of the total Peruvian booby ASVs (61.3%) than of the total Guanay cormorant ASVs (33.1%).

Influence of physicochemical parameters on the microbiota associated with island guano

Correlations among guano physicochemical parameters were assessed using a Spearman correlation matrix, including pH, electrical conductivity (EC), total nitrogen (N%), organic matter (OM%), organic carbon (OC%), C/N ratio, phosphorus (P_2O_5%), potassium (K_2O%), iron (Fe), copper (Cu), zinc (Zn), and manganese (Mn) (Fig. 8).

The strongest positive correlation observed was between OM% and OC%, with a coefficient of r = 1.00, indicating a perfect correspondence between these two variables. Similarly, OM% and OC% showed strong, statistically significant negative correlations with K_2O% ($r = -0.77$, $p < 0.01$ in both cases) and with Mn ($r = -0.66$, $p < 0.01$ in both cases). The C/N ratio showed a significant negative correlation with N% ($r = -0.67$, $p < 0.05$). pH exhibited a strong and significant negative correlation with Mn ($r = -0.70$, $p < 0.01$), a moderate negative correlation with Fe ($r = -0.59$) and with K_2O% ($r = -0.47$), and a moderate positive correlation with OM% and OC% (r = 0.45 in both cases). EC showed a significant positive correlation with K_2O% ($r = 0.64$, $p < 0.05$) and a significant negative correlation with Cu ($r = -0.60$, $p < 0.05$). K_2O% in turn showed a significant negative correlation with Cu ($r = -0.66$, $p < 0.05$) and moderate positive correlations with P_2O_5% (r = 0.51) and Mn (r = 0.51). Fe showed a moderate positive correlation with Mn (r = 0.56). Correlations between P_2O_5% and the remaining variables were generally weak, with the exception of its moderate positive correlation with K_2O% (r = 0.51) and its moderate negative correlation with Cu ($r = -0.42$). Zn and Cu showed weak correlations with most of the parameters evaluated. N% showed moderate positive correlations with OM% and OC% (r = 0.46 in both cases) and a weak negative correlation with EC ($r = -0.20$).

The relationship between microbial community composition and guano physicochemical variables was assessed using redundancy analysis (RDA) for the 15 most abundant bacterial and fungal genera (Fig. 9).

The bacterial RDA explained 39.7% of variance on the RDA1 axis and 35.7% on the RDA2 axis, accounting for a cumulative 75.4% of total constrained variance across the first two canonical axes (Fig. 9a).

Among the physicochemical variables, K_2O%, Fe, and Zn showed a positive association with one another and were opposed to OM% and OC%, which showed a negative association relative to these three parameters. pH was positively associated with Zn and negatively associated with OM% and OC%. EC

was positively associated with K_2O% and Fe. Mn and P_2O_5% showed an opposing association to K_2O% and Fe. N% and C/N were negatively associated with each other along the main gradient.

Regarding sample distribution, Guañape Norte (Booby) was positively associated with K_2O% and Fe, representing the sample with the greatest affinity toward these parameters. Punta Coles (Guanay) and Chincha Norte (Booby) were positively associated with P_2O_5% and Mn. Isla Pescadores (Guanay) was positively associated with OM% and OC%. Isla Santa Rosa (Guanay) and Isla Macabí (Guanay) were positively associated with pH and Zn. The majority of the remaining samples showed no marked association with any particular variable.

Regarding bacterial genera, *Paeniclostridium* was positively associated with K_2O% and Fe. *Gottschalkia* was positively associated with OM% and OC% and negatively associated with K_2O% and Fe. *Atopostipes* was positively associated with pH and negatively associated with OM% and OC%. *Bacillus*, *Neobacillus*, and *Pseudogracilibacillus* were positively associated with pH and Zn. *Clostridium*, *Saccharomonospora*, and *Pseudomonas* showed a positive association with EC. *Lentibacillus*, *Kushneria*, and *Virgibacillus* showed no marked association with any particular variable. *Ammoniphilus* was positively associated with OM% and OC% and negatively associated with K_2O%.

The fungal RDA explained 57.3% of variance on the RDA1 axis and 25.5% on the RDA2 axis, accounting for a cumulative 82.8% of total constrained variance across the first two canonical axes (Fig. 9b).

Among the physicochemical variables, K_2O%, P_2O_5%, Zn, Fe, and Mn showed a positive association with one another and were opposed to EC, which showed a negative association relative to these parameters. OM%, OC%, C/N, and pH were positively associated with one another and negatively associated with K_2O%, P_2O_5%, Zn, Fe, and Mn. N% was negatively associated with OM% and OC%.

Regarding sample distribution, Guañape Norte (Booby) showed a strong positive association with EC, being the sample most distant from the rest in ordination space. Isla Macabí (Guanay) and Chincha Norte (Guanay) were positively associated with OM%, OC%, C/N, and pH. Isla Santa Rosa (Guanay) and Chincha Norte (Booby) likewise showed a positive association with OM% and OC%. Guañape Norte (Guanay), Isla Santa (Guanay), Guañape Sur (Booby), and Isla Pescadores (Guanay) samples were positively associated with K_2O%, Fe, Zn, and P_2O_5%.

Regarding fungal genera, *Debaryomyces* was positively associated with pH and C/N and negatively associated with K_2O% and EC. *Penicillium* was positively associated with EC and negatively associated with K_2O% and P_2O_5%. *Emmonsiellopsis* was positively associated with EC and negatively associated with OM% and OC%. *Aspergillus* was positively associated with K_2O% and Mn and negatively associated with OM%, OC%, and pH. *Leucothecium* was positively associated with Mn and Fe. *Mortierella*, *Spiromastigoides*, *Rhizophlyctis*, and Fungi_gen_Incertae_sedis were positively associated with EC and negatively associated with K_2O% and P_2O_5%. *Ascobolus* was positively associated with OM%, OC%, and EC. *Alternaria*, *Preussia*, *Chrysosporium*, *Talaromyces*, and *Acrophialophora* showed no marked association with any particular variable, positioning around the ordination origin.

The functional redundancy of bacterial and fungal communities was estimated using FAPROTAX and FUNGuild, respectively, and their relationship with guano physicochemical variables was assessed using Spearman correlations (Fig. 10).

Bacterial functional redundancy analysis identified a broad repertoire of predicted biogeochemical functions in both bird species (Fig. 10a). Aerobic chemoheterotrophy and anaerobic chemoheterotrophy were the functions with the highest functional redundancy in both groups, with values exceeding 30% of weighted abundance. Other functions with considerable representation included fermentation, nitrogen respiration, nitrate reduction, and methylotrophy. Functions associated with organic compound degradation — including aromatic hydrocarbon degradation, sulfur compound degradation, and manganese degradation — showed lower but detectable redundancy in both groups. Differences in functional redundancy between Guanay cormorant and Peruvian booby were visually small across most functional categories, with overlapping standard error bars between groups. Functions with the greatest between-group variability included methylotrophy, nitrate reduction, and nitrification.

The distribution of bacterial biogeochemical functions by locality showed variable functional composition across sites (Fig. 10e). Guañape Norte (Booby) and Guañape Sur (Booby) samples presented visually distinct functional profiles relative to the remaining localities, with a higher relative proportion of certain functional categories. Isla Pescadores (Guanay), Isla Santa Rosa (Guanay), and Isla Mazorca (Guanay) samples showed more homogeneous functional profiles among themselves.

The correlation heatmap between bacterial functions and physicochemical variables showed associations of varying direction and magnitude depending on the function evaluated (Fig. 10c). Chemoheterotrophic functions showed positive correlations with EC and negative correlations with OM% and OC%. Sulfur compound degradation showed negative correlations with N% and C/N. Methylotrophy showed positive correlations with K_2O% and negative correlations with OM% and OC%. Functions associated with the nitrogen cycle — including nitrification, nitrate reduction, and nitrogen respiration — showed variable correlation patterns depending on the physicochemical variable evaluated. Statistically significant correlations ($p_adj < 0.05$ and $p < 0.10$) corresponded primarily to functions with the highest functional redundancy and to the variables EC, K_2O%, OM%, OC%, and N%.

Fungal ecological guild analysis identified saprotrophs (Saprotroph) as the absolutely dominant guild in both bird species, with a mean functional redundancy exceeding 75% of weighted abundance in both groups (Fig. 10b). Wood saprotrophs (Wood Saprotroph) constituted the second most represented guild, with variable values between Guanay cormorant and Peruvian booby. Pathotrophic fungi (Pathotroph) represented the third most abundant guild, followed by symbiotrophs (Symbiotroph), endophytes (Endophyte), and animal pathogens (Animal Pathogen). Guilds with lower representation included soil saprotrophs (Soil Saprotroph), dung saprotrophs (Dung Saprotroph), plant pathogens (Plant Pathogen), epiphytes (Epiphyte), litter saprotrophs (Litter Saprotroph), fungal parasites (Fungal Parasite), bryophyte parasites (Bryophyte Parasite), and ectomycorrhizal fungi (Ectomycorrhizal). Differences in functional

redundancy between Guanay cormorant and Peruvian booby were visually small across most guilds, with overlapping standard error bars.

The distribution of fungal ecological guilds by locality showed a saprotroph-dominated composition across all samples evaluated, with variation in the relative proportion of secondary guilds among sites (Fig. 10f). Guañape Norte (Booby) and Guañape Sur (Booby) samples presented visually distinct guild profiles relative to the remaining samples, with a lower proportion of saprotrophs and greater relative representation of other guilds. The Isla Pescadores (Guanay) sample showed a distinctive guild profile, consistent with its peripheral position observed in beta diversity analyses.

The correlation heatmap between fungal guilds and physicochemical variables showed variable associations depending on the guild evaluated (Fig. 10d). Wood saprotrophs showed negative correlations with EC and K_2O%, and positive correlations with OM%, OC%, and C/N. Soil saprotrophs showed negative correlations with EC and K_2O%. Plant pathogens showed positive correlations with N% and negative correlations with some micronutrients. Endophytes and epiphytes showed positive correlations with Fe and Zn. Animal pathogens showed negative correlations with OM% and OC%. Statistically significant correlations ($p_{adj} < 0.05$ and $p < 0.10$) were sparse, corresponding primarily to the most represented guilds and to the variables EC, K_2O%, OM%, and OC%.

Spearman correlations between the most abundant bacterial and fungal genera and guano physicochemical variables were assessed with false discovery rate (FDR) correction (Fig. 11). Spearman correlation analysis between dominant bacterial genera and physicochemical variables revealed associations of varying direction and magnitude (Fig. 11a). *Saccharomonospora* showed the strongest and most statistically significant correlations in the dataset, with positive correlations with K_2O% ($\rho = 0.80$, $p < 0.05$), Fe ($\rho = 0.66$, $p < 0.10$), and EC ($\rho = 0.54$), and significant negative correlations with OC% ($\rho = -0.79$, $p < 0.05$) and OM% ($\rho = -0.79$, $p < 0.05$). *Clostridium* likewise showed significant negative correlations with OC% ($\rho = -0.79$, $p < 0.05$) and OM% ($\rho = -0.79$, $p < 0.05$), and positive correlations with EC ($\rho = 0.54$) and Fe ($\rho = 0.46$). *Neobacillus* showed significant positive correlations with OC% ($\rho = 0.72$, $p < 0.05$) and OM% ($\rho = 0.72$, $p < 0.05$), and a negative correlation with Mn ($\rho = -0.45$). *Bacillus* showed a positive correlation with EC ($\rho = 0.67$) and moderate negative correlations with OC% and OM% ($\rho = -0.16$ in both cases). *Pseudomonas* showed negative correlations with EC ($\rho = -0.66$, $p < 0.10$) and K_2O% ($\rho = -0.53$). *Tyzzerella* showed a negative correlation with N% ($\rho = -0.56$) and positive correlations with P_2O_5% ($\rho = 0.47$) and OC% and OM% ($\rho = 0.46$ in both cases). *Ammoniphilus* showed a negative correlation with Mn ($\rho = -0.63$) and moderate negative correlations with K_2O% ($\rho = -0.38$) and N% ($\rho = -0.31$). *Gottschalkia* showed a positive correlation with Zn ($\rho = 0.60$) and a negative correlation with Mn ($\rho = -0.46$). *Lentibacillus* showed negative correlations with Fe ($\rho = -0.41$), K_2O% ($\rho = -0.37$), and Zn ($\rho = -0.58$). *Oceanobacillus* showed a positive correlation with EC ($\rho = 0.50$). *Virgibacillus*, *Paeniclostridium*, *Kushneria*, *Pseudogracilibacillus*, and *Atopostipes* showed low-magnitude correlations with most variables evaluated, reaching no adjusted statistical significance.

Spearman correlation analysis between dominant fungal genera and physicochemical variables revealed generally weaker association patterns than those observed for bacteria, with fewer statistically significant correlations (Fig. 11b). *Aspergillus* showed the strongest correlations in the fungal dataset, with significant positive correlations with K_2O% (ρ = 0.69, p < 0.10) and Mn (ρ = 0.66, p < 0.10), and significant negative correlations with OC% (ρ = −0.59, p < 0.10), OM% (ρ = −0.59, p < 0.10), and pH (ρ = −0.50). *Penicillium* showed a positive correlation with EC (ρ = 0.68) without reaching adjusted significance. Fungi_gen_Incertae_sedis showed a positive correlation with EC (ρ = 0.59) and a negative correlation with N% (ρ = −0.39). *Emmonsiellopsis* showed positive correlations with EC (ρ = 0.47) and Mn (ρ = 0.42), and a negative correlation with pH (ρ = −0.49). *Spiromastigoides* showed a positive correlation with Mn (ρ = 0.52) and negative correlations with C/N (ρ = −0.51) and Zn (ρ = −0.50). *Talaromyces* showed negative correlations with C/N (ρ = −0.44), OC% (ρ = −0.41), and OM% (ρ = −0.41), and a positive correlation with Mn (ρ = 0.44). *Debaryomyces* showed a negative correlation with K_2O% (ρ = −0.47) and positive correlations with OC% (ρ = 0.30) and OM% (ρ = 0.30). *Leucothecium* showed negative correlations with OC% (ρ = −0.43) and OM% (ρ = −0.43). *Alternaria*, *Ascobolus*, *Chrysosporium*, *Mortierella*, *Preussia*, *Rhizophlyctis*, and *Acrophialophora* showed low to moderate correlation magnitudes with the physicochemical variables evaluated, reaching no adjusted statistical significance in any case.

Discussion

The present study constitutes the first comprehensive, culture-independent characterization of bacterial and fungal communities associated with the guano of the principal seabirds of the Peruvian coast, *Leucocarbo bougainvillii* and *Sula variegata*, using high-throughput amplicon metabarcoding of the 16S rRNA and ITS markers. The complete absence of prior studies on the guano microbiome of these species, in contrast to the body of research available on Antarctic penguin guano [13,34] and cave-dwelling bat guano [16,18,20], underscores the pioneering nature of this work and establishes a fundamental microbiological baseline for the guano island ecosystem of the South Pacific. The integration of taxonomic, ecological, and functional data across 13 guano zones distributed along more than 2,000 km of coastline further allows interpretation of microbial patterns in a biogeographic and agronomic context of strategic relevance to Peru.

The near-absolute dominance of Bacillota, with relative abundances exceeding 90% across all samples, represents the most consistent and ecologically informative taxonomic finding of the study. This pattern has no direct precedent in the seabird guano literature, although it bears some correspondence with studies on Adélie penguin guano reporting Firmicutes as one of the dominant phyla in mature deposits alongside Proteobacteria and Bacteroidetes [13]. However, the proportion observed in the present study is considerably higher than that reported in any previously characterized guano system, suggesting that the specific physicochemical conditions of Peruvian island guano, characterized by an extraordinary concentration of nitrogenous organic matter and variable salinity, exert an exceptionally selective environmental filter on the bacterial community. The preponderance of the order Bacillales and the

family Bacillaceae is consistent with the known ecology of this group, which encompasses endospore-forming capacity, desiccation tolerance, resistance to osmotic fluctuations, and metabolic versatility in organic matter-rich substrates [35]. The presence of halotolerant genera such as *Virgibacillus* and *Lentibacillus* reinforces this pattern: *Virgibacillus* possesses well-documented molecular halotolerance mechanisms including ectoine synthesis and ion channel regulation [36], while *Lentibacillus* has been characterized as a moderate halophile with specific genomic adaptations for high-salinity environments [37]. The coexistence of stress-resistant endospore-forming bacteria with strict anaerobic fermenters such as *Gottschalkia*, *Atopostipes*, and *Tissierella* reflects the internal redox heterogeneity of the guano substrate, where the progressive accumulation of organic matter generates anoxic microenvironments sustaining differentiated metabolic consortia, a pattern analogous to that described in vertically stratified bat guano [18].

Within the exploratory scope of this study, no clear host-associated pattern emerged in alpha or beta diversity, suggesting that substrate characteristics may be at least as important as bird identity in structuring the guano microbiome. This result indicates that both organic matrices sustain bacterial communities of comparable richness and evenness, regardless of differences in diet, metabolism, and physiology between the two bird species. Similar patterns of bacterial diversity homogeneity across hosts have been documented in comparative studies of guano from different bat species, where substrate composition and environmental conditions at the deposition site tend to override the phylogenetic influence of the host in structuring bacterial communities [15,16]. In systems with high carbon and nitrogen availability, such as island guano, ecological theory predicts the formation of functionally redundant communities in which distinct taxa share equivalent metabolic capacities [26,38], which would explain the stability in diversity indices observed between groups. Beta diversity analyses using PERMANOVA confirmed this pattern: none of the dissimilarity metrics evaluated, Bray–Curtis or Jaccard, detected significant differences in the overall bacterial community composition between Guanay cormorant and Peruvian booby, suggesting that bird species identity does not predominantly determine the structure of the guano bacterial microbiome. This result contrasts with studies in penguins where the host species or individual age generates detectable differences in the gastrointestinal microbiota [39,40], although it should be noted that those studies analyzed intestinal contents rather than deposited guano, which represents a subsequent stage of microbial transformation with its own dynamics.

The results of the redundancy analysis suggest that the physicochemical gradient of the guano, rather than host identity, constitutes the principal structuring axis of microbial communities in these island ecosystems. The RDA model explained 75.4% of constrained bacterial variance and 82.8% of fungal variance, proportions considerably higher than those reported in soil microbiome studies where environmental gradients typically explain between 20% and 50% of compositional variance [29,41]. This high proportion of variance explained by physicochemical variables suggests that island guano represents an especially intense environmental filtering system, consistent with the hypothesis of pH-mediated deterministic assembly proposed by Tripathi [42]. The specific associations identified in the bacterial RDA, halotolerant genera such as *Paeniclostridium* and *Saccharomonospora* positively

associated with K_2O% and Fe, and fermenters such as *Gottschalkia* and *Ammoniphilus* associated with OM% and OC%, are consistent with the known physiological preferences of these groups. Salinity, measured as electrical conductivity, has been identified as one of the most powerful environmental determinants of microbial community composition at a global scale [43–45], and the results of the present study are fully consistent with this theoretical framework. Spearman correlations reinforced this pattern by identifying *Saccharomonospora* as the genus with the strongest and most statistically significant physicochemical correlations in the bacterial dataset, with positive associations with K_2O% ($\rho = 0.80$, $p < 0.05$) and negative associations with OM% ($\rho = -0.79$, $p < 0.05$), positioning it as a potential candidate bioindicator of the chemical maturation of the guano substrate.

Fungal communities displayed an even more simplified structure than bacterial ones, with Ascomycota dominance exceeding 95% across all samples and near-exclusive concentration of diversity within the order Eurotiales and the family Aspergillaceae. This structure is consistent with global patterns of Ascomycota dominance in readily available organic matter-rich substrates, where a small number of generalist, stress-tolerant phylotypes tend to monopolize resources [46,47]. The presence of *Aspergillus* as the dominant genus and the only one with significant positive differential-abundance signal in Guanay cormorant relative to Peruvian booby is consistent with the role of this genus as a ubiquitous saprotroph with high enzymatic versatility and elevated osmotic tolerance [48]. Studies of bat guano in Bornean caves likewise documented *Aspergillus* as the most frequent fungal genus, representing up to 36.4% of guano isolates [49], suggesting that the dominance of this genus in organically rich fecal substrates may represent a transcontinental, host-independent pattern. The presence of *Chrysosporium*, a keratinolytic genus with demonstrated capacity to degrade feathers, keratin, and collagen as the sole source of C, N, and S [50,51], is ecologically consistent with a seabird-derived substrate that inevitably contains plumage material in various stages of decomposition. *Mortierella*, detected at relevant abundances in some northern localities, is a lipid-producing saprotroph documented as a plant growth promoter in agricultural soils [52], whose distribution in guano was significantly influenced by pH [53], in line with its positive association with OM% and C/N observed in the fungal RDA of the present study. Furthermore, the presence of strict coprophilous fungi such as *Preussia* and *Ascobolus*, systematically documented in vertebrate excreta [54,55], confirms the specificity of the guano ecological niche as a fecal substrate and validates the quality of the sampling conducted.

Functional analysis predicted using FAPROTAX and FUNGuild suggested a broadly similar predicted metabolic profile between Guanay cormorant and Peruvian booby guano, with no significant differences in any of the functional categories evaluated following correction for multiple comparisons.Given that these functions were inferred bioinformatically and that the study design was exploratory, the observed functional stability should be interpreted as a preliminary pattern requiring confirmation through replicated sampling and, ideally, direct functional assays. The dominance of aerobic and anaerobic chemoheterotrophy, fermentation, and nitrogen respiration functions is consistent with the central role of these pathways in the mineralization of concentrated organic matter and in the biogeochemical transformation of nitrogenous compounds, particularly uric acid and its derivatives, that characterize

guano as a substrate [17]. This functional stability among communities taxonomically structured by environmental gradients is consistent with the concept of microbial functional redundancy, whereby distinct taxa can sustain equivalent processes under similar selective pressures [27,38,56]. At the fungal level, the dominance of the saprotrophic guild (> 75%) estimated by FUNGuild confirms that the guano mycobiome is organized primarily around the decomposition of complex organic matter, with a notable fraction of pathotrophs that may include opportunistic phytopathogenic fungi [57]. The stability of functional profiles despite the taxonomic variation observed among localities reinforces the notion that Peruvian island guano maintains a robust and predictable metabolic machinery, with direct implications for its valorization as a biofertilizer resource: regardless of the extraction locality or the producing bird species, guano would contribute a relatively homogeneous microbial functional profile oriented toward nutrient mineralization and soil fertility improvement [10].

Taken together, the results of the present study establish that the Peruvian island guano microbiome is a chemically filtered, taxonomically simplified yet functionally stable microbial system, whose structure is governed primarily by the abiotic gradients of the substrate, particularly electrical conductivity, pH, and organic matter content, rather than by the identity of the host bird species. This inaugural characterization opens multiple avenues for future research, including longitudinal assessment of microbial dynamics during guano extraction and storage, identification of strains with specific plant growth-promoting and nitrogen-fixing capacities, and development of microbiological monitoring protocols for the sustainable management of guano as a biofertilizer in Peruvian agroecosystems. The integration of metabarcoding as a routine guano quality control tool would, in perspective, enable the selection of batches with optimal microbial profiles for specific agricultural applications, maximizing the ecosystem services of this strategic resource of the Peruvian coastline.

Conclusions

The present study provides the first culture-independent baseline description of bacterial and fungal communities associated with the guano of the principal guano-producing seabirds of the Peruvian coast, *Leucocarbo bougainvillii* and *Sula variegata*, revealing a taxonomically simplified yet functionally robust microbiome dominated by Bacillota (> 90%) at the bacterial level and by Ascomycota (> 95%) at the fungal level, with halotolerant and endospore-forming genera as central components across all guano zones evaluated. The absence of significant differences in alpha and beta diversity between both bird species, combined with the high proportion of microbial variance explained by the physicochemical variables of the substrate, 75.4% for bacteria and 82.8% for fungi, shows that the chemical gradient of the guano, particularly electrical conductivity, pH, and organic matter content, may be an important ecological filter of microbiome structure. Predicted functional profiles using FAPROTAX and FUNGuild were broadly similar across samples, suggesting potential functional stability, although this inference remains provisional, with predominance of chemoheterotrophic, fermentative, nitrogen mineralization, and saprotrophic functions, a pattern consistent with microbial functional redundancy that positions Peruvian island guano as a resource with a predictable biogeochemical potential independent of its

geographic or avian origin. These findings provide the scientific basis for the incorporation of metabarcoding as a microbiological monitoring tool in the sustainable management of guano, opening concrete prospects for the selection of batches with optimal microbial profiles destined for the biofertilization of degraded Peruvian agroecosystems.

Methods

Study area and sampling design

The study was conducted in seabird guano harvesting zones along the Peruvian coastline. Ten geographic study units were selected, corresponding to eight islands and two coastal guano headlands, based on: (i) their level of guano deposition activity, (ii) their suitability for extraction, and (iii) their relevance within the management system administered by the Guano Extraction and Commercialization Sub-unit of the Rural Productive Development Program (AGRORURAL), under the Ministry of Agrarian Development and Irrigation of Peru. As these areas form part of protected natural reserves, access was granted following authorization from the National Service of Natural Protected Areas (SERNANP), an agency under the Ministry of the Environment.

Sampling zones were distributed across different biogeographic regions of the Peruvian coast. On the southern coast, sites were included in Moquegua (one headland) and Ica (two islands and one headland); on the central coast, in Lima (two islands) and Ancash (one island); and on the northern coast, in La Libertad (three islands) (Fig. 1; Table S1). Sampling focused on the two principal guano-producing seabird species of the Peruvian coast: the Guanay cormorant (*Leucocarbo bougainvillii*) and the Peruvian booby (*Sula variegata*). At sites where deposits associated with both species were identified, each species was treated separately as an independent sampling unit. Consequently, although ten geographic sites were evaluated, 13 site–species sampling units were obtained, yielding 13 composite samples. Sample G4 (Isla Lobos de Tierra) was excluded from all analyses due to insufficient metadata regarding the dominant bird species and uncertainty about the nature of the collected substrate (guano vs. soil).

Fieldwork was conducted between November 11 and December 27, 2024. At each site–species sampling unit, 20 spatially distributed subsamples were collected from within active deposition areas and subsequently homogenized to generate a single representative composite sample. In total, 260 georeferenced subsamples and 13 composite samples were obtained. Sample identifiers, geographic provenance, and associated bird species are detailed in Table S1.

As the objective of this study was to characterize spatial and ecological patterns of the guano microbiome at the site scale, subsamples from each unit were integrated into a single composite sample. Because each site–bird species unit was represented by a single composite sample, this study lacks within-site biological replication. Consequently, all between-group statistical analyses, including alpha diversity, beta diversity, differential abundance, and ordination-based associations, were

interpreted as exploratory pattern-detection analyses at the composite-sample level rather than as confirmatory tests of bird-species effects.

Sample collection and preservation

Guano samples were collected from active surface deposits at each study site. Recent deposition activity was verified by direct observation and confirmed with information provided by AGRORURAL island wardens, who also identified the predominant bird species at each zone.

At each site, 20 simple sampling points were established following a zigzag design to maximize spatial coverage. Sampling was conducted using personal protective equipment and sterile materials, including nitrile gloves, face masks, boots, 90% alcohol, disinfected excavation tools, a 10 × 10 cm sampling frame, a portable scale, and a global positioning system (GPS) unit.

At each sampling point, the uppermost guano layer was removed to eliminate potentially contaminating material such as salt crusts, feathers, and other organic debris. The deposit was then excavated to a maximum depth of 30 cm, exposing a vertical face from which a block of approximately 5 cm thickness was extracted. The outer edges of each block were discarded to minimize contamination, and each subsample was georeferenced using GPS.

A total of 20 subsamples were collected per site, each weighing approximately 850 g and covering a depth range of 0–30 cm. Subsamples were transported in clean containers and processed on a previously disinfected plastic surface, where visible impurities — including shell fragments, stones, ectoparasites, feathers, and bone remains — were removed.

The material was thoroughly homogenized and sieved through a 2 mm mesh to obtain a representative composite sample. Aliquots were then taken for the respective analyses. For amplicon metabarcoding analysis, approximately 50 g of island guano per sample was collected from the homogenized material and divided into three independent portions. One portion was preserved in sterile microtubes containing DNA/RNA Shield (Zymo Research, USA) at a ratio of 100 mg of sample per 800 µL of reagent. The remaining portions were stored in sterile 50 mL tubes and in single-use sampling bags.

All samples were handled using sterile materials and transported under cold chain conditions in insulated boxes with ice packs. Upon arrival at the laboratory, samples were stored at − 80°C and − 40°C until processing for DNA extraction.

Physical and chemical analysis

The physical and chemical properties of each composite sample were determined using accredited methods under the ISO/IEC 17025 technical standard at the Soil, Water, and Foliar Analysis Laboratory (LABSAF) of the Canaán Agricultural Experimental Station of the National Institute of Agrarian Innovation (INIA). The variables analyzed included pH (in 1:2.5 aqueous suspension), electrical conductivity (EC, dS m^{-1}), organic matter percentage (%OM), total organic carbon (%OC), total nitrogen (%N), phosphorus expressed as $\%P_2O_5$, total potassium expressed as $\%K_2O$, and micronutrients including Fe, Cu, Zn, and

Mn, expressed in ppm. These parameters were selected based on their well-recognized ecological relevance in structuring microbial communities in soils and organic substrates [29,30].

Prior to multivariate analysis, collinearity among physicochemical variables was assessed using a Spearman correlation matrix. The variance inflation factor (VIF) was calculated for all remaining variables, and only those with VIF < 10 were retained to avoid severe multicollinearity in the constrained ordination. Based on this procedure, the following eleven physicochemical variables were included as environmental predictors in the RDA models: pH, electrical conductivity (EC), total nitrogen (N%), organic matter (OM%), C/N ratio, phosphorus (P_2O_5%), potassium (K_2O%), iron (Fe), copper (Cu), zinc (Zn), and manganese (Mn).

DNA Extraction and Amplicon Sequencing

Total genomic DNA was extracted from 200 mg of each island guano sample (Table S1) using the E.Z.N.A.® Soil DNA Kit (Cat. No. D5625; Omega Bio-tek, Norcross, USA). The manufacturer's protocol for 250 mg to 1 g samples was followed, with a modification to the mechanical lysis step to increase disruption efficiency. Guano samples (200 mg) were transferred to bead-beating tubes containing glass beads and subjected to horizontal mechanical agitation at 1,500 rpm for 15 minutes using a Vortex Genie 2 Digital mixer (Scientific Industries, USA). The remainder of the protocol was then carried out according to the manufacturer's instructions. DNA was eluted in 50 µL of elution buffer pre-warmed to 70°C. The concentration and purity of extracted DNA were assessed by, UV spectrophotometry using a Synergy H1 microplate reader (Agilent Technologies, USA), and 0.8% agarose gel electrophoresis. Samples meeting high-quality criteria for integrity and purity were stored at − 20°C until shipment to a commercial genomics facility for sequencing on the Illumina MiSeq platform.

Sequencing was performed at a commercial genomics facility (China) on the Illumina MiSeq system in paired-end mode (2 × 150 bp). For bacterial community analysis, the V3–V4 hypervariable region of the 16S rRNA gene was amplified using primers 341F/806R [58]. For fungal communities, the ITS1–ITS2 region was amplified using primers ITS1 (5'-TCCGTAGGTGAACCTGCGG-3') and ITS2 (5'-GCTGCGTTCTTCATCGATGC-3') [59,60]. Both libraries were sequenced independently.

Bioinformatic Analysis

Bioinformatic processing was performed using the *nf-core/ampliseq* pipeline [61], executed with Nextflow v4.2.1 [62] within a Singularity container. Reproducible software environments were provided through Bioconda [63] and BioContainers [64]. Raw sequence quality was assessed with FastQC [65] and consolidated reports were generated with MultiQC [66].

Sequence processing was carried out using theDADA2 v1.26.0 package [67], encompassing demultiplexing, quality-based truncation, paired-end read merging, chimera removal, and amplicon sequence variant (ASV) inference. Taxonomic assignment was performed within DADA2 using a minimum bootstrap confidence threshold of 50 (based on 100 iterations). The SILVA v138 database [68]

was used as reference for bacterial sequences (16S marker), and the UNITE v8.3 database [69]. ASVs together with their abundance counts and taxonomic annotations were imported into QIIME2 v2023 [70]. ASVs identified as mitochondria or chloroplasts were removed prior to downstream analyses.

All statistical analyses and data visualization were conducted in R v4.4.2. The packages and procedures used at each analytical stage are described below. Phyloseq objects generated by the bioinformatic pipeline were imported and manipulated using the *phyloseq* v1.42.0 package [71]. Sequences assigned as mitochondria, chloroplasts, or lacking phylum-level assignment were removed prior to all analyses. Two types of transformation were applied to the ASV tables: rarefaction to the minimum sequencing depth per marker using the rarefy_even_depth function with a fixed random seed (set.seed = 123) for alpha diversity analyses, and normalization to relative abundance for compositional analyses, beta diversity, and ordination.

Alpha diversity indices — Observed features, Chao1, Shannon, and Simpson — were calculated using the *estimate_richness* function in phyloseq. Differences in alpha diversity indices between Guanay cormorant and Peruvian booby guano were assessed using the non-parametric Kruskal–Wallis test, implemented with the *kruskal.test* function from base R. Visualization was performed using violin and box plots generated with *ggplot2* v3.5.0 [72] and *ggpubr* [73], with individual sample labeling via *ggrepel* [74].

Beta diversity was evaluated through Principal Coordinates Analysis (PCoA) and Non-Metric Multidimensional Scaling (NMDS), using Bray–Curtis and Jaccard dissimilarity metrics — the latter in binary mode — calculated with the distance function in phyloseq. Ordination was performed using the ordinate function in phyloseq and the rda function in vegan. Significant differences in community composition between groups were assessed using Permutational Multivariate Analysis of Variance (PERMANOVA) with the adonis2 function in vegan (999 permutations), with bird species and geographic region as explanatory variables. Dispersion homogeneity between groups was verified using the *betadisper* function in vegan, followed by a permutation-based ANOVA. 95% confidence ellipses were overlaid on PCoA biplots when a group contained three or more samples.

Stacked bar plots of relative abundance were generated at the phylum, order, family, and genus levels using the tax_glom and psmelt functions in phyloseq, selecting the most abundant taxa at each level (up to 20 genera). Visualization was performed with ggplot2 using color palettes from RColorBrewer [75]. Euler diagrams to visualize ASVs shared and exclusive between Guanay cormorant and Peruvian booby were constructed using the *eulerr* v7.0.0 package [76], with circles proportional to the number of ASVs in each group. Differential abundance analysis between Guanay cormorant and Peruvian booby guano was performed using the DESeq2 method, implemented with the *DESeq2* v1.42.0 package [77], applying the Wald test with library size estimation by the positive counts (poscounts) method. ASV counts were rounded to integers prior to analysis. Results were filtered at an adjusted significance threshold of $p < 0.05$ with a Guanay vs. Booby contrast. Volcano plots and $\log_2$ fold-change plots were generated with *ggplot2* and *ggrepel*.

Correlations among guano physicochemical parameters were assessed using Spearman's correlation coefficient with the *cor* function from base R, and visualized as correlation matrices generated with the *corrplot* v0.95 package [78]. Statistical significance for each pairwise correlation was evaluated using cor.mtest, with significance levels $p < 0.05$, $p < 0.01$, and $p < 0.001$ indicated directly within each matrix cell.

The relationship between microbial community composition and guano physicochemical variables was assessed using redundancy analysis (RDA) with the rda function in vegan, following Hellinger transformation of the ASV table using the decostand function. pH, EC, N%, OM%, OC%, C/N ratio, P_2O_5%, K_2O%, Fe, Zn, and Mn were included as predictor variables. The overall significance of the model and of each individual variable was evaluated using permutation tests with the anova.cca function (999 permutations). The adjusted coefficient of determination (adjusted R^2) was calculated using the RsquareAdj function. For the genus-level biplot, analysis was performed on the 15 most abundant genera per marker, excluding unclassified genera. Biplots were generated with ggplot2 and ggrepel.

Spearman correlations between the relative abundance of dominant genera and physicochemical variables were calculated individually for each taxon–variable pair using the cor.test function from base R. P-values were corrected for the false discovery rate (FDR) using the Benjamini–Hochberg method with the p.adjust function. Results were visualized as heatmaps generated with ggplot2, with correlation coefficients and significance levels ($p_{adj} < 0.05$ and $p_{adj} < 0.10$) overlaid on each cell.

The functional profile of bacterial communities was estimated using the FAPROTAX database, implemented through the *microeco* package [79] with the *trans_func* class. Functional redundancy was calculated with abundance weighting (abundance_weighted = TRUE), expressed as a percentage of the total community. Differences in functional redundancy between Guanay cormorant and Peruvian booby guano were assessed using the Kruskal–Wallis test. Visualizations included heatmaps generated with *pheatmap* [80] and grouped bar plots with ggplot2 and patchwork [81].

Ecological guilds of fungal communities were assigned using the FUNGuild [57] and FungalTraits databases, also implemented through microeco with the trans_func class, including assignments at confidence levels "Highly Probable," "Probable," and "Possible." Functional redundancy was calculated with abundance weighting, and differences between groups were assessed using the Kruskal–Wallis test. Spearman correlations between estimated functional redundancy — from both FAPROTAX and FUNGuild — and physicochemical variables were calculated and visualized following the same protocol described for microbial genera.

Declarations

Data availability

The sequencing data generated in this study are publicly available in the NCBI BioProject repository under accession number PRJNA1454567. All other data generated or analyzed during this study are included in this published article and its Supplementary Information files.

Competing interests

The authors declare no competing interests.

Consent for publication

Each participant provided explicit written informed consent authorizing the use and publication of their photographic images in the context of this scientific article and in any accompanying materials, including online versions. All participants were informed about the scope of dissemination, including open-access publication, and acknowledged that their images may be publicly available.

Additional information

All additional information is included in Supplementary Materials.

Funding

This work was funded by CONCYTEC through the PROCIENCIA program within the framework of the E041-2023-04 call for AECID Scientific Research Projects to reduce vulnerability to the effects of the climate, food, and/or health crisis in Peru from the IPIs 2023-04, under contract PE501086806-2024-PROCIENCIA, in the framework of Grant Agreement Exp. No. 2021/SPE/0000400012 with AECID.

Author Contribution

R.C. conceptualized the study, designed the methodology, collected guano samples, conducted the investigation, curated the data, wrote the original draft, created visualizations, acquired funding, and supervised the project. H.C.-S. performed the formal analysis and statistical analysis of all data, created figures and tables, curated the data, wrote the original draft, and reviewed and edited the manuscript. E.M. reviewed and edited the manuscript. An.C. conducted the investigation, collected samples, performed laboratory processing, and curated the data. M.P.P. conducted the investigation, developed the DNA extraction methodology, and curated the data. V.R. conducted the investigation, performed formal analysis, created visualizations, and reviewed and edited the manuscript. L.P. developed the sampling methodology and protocol and conducted the investigation. L.S. developed the DNA extraction methodology, wrote the methodology section of the manuscript, and reviewed and edited the manuscript. Ar.C. administered the project, provided resources, and reviewed and edited the manuscript. R.V. conducted the investigation, provided advisory support on DNA extraction methodology, and reviewed and edited the manuscript. S.M. reviewed and edited the manuscript. All authors read and approved the final manuscript.

Acknowledgement

The authors wish to thank the Directorate for the Management of Protected Natural Areas of the National Service of Protected Natural Areas by the State (SERNANP), under the Ministry of the Environment of Peru, for granting the necessary permits to access the Guano Islands and Headlands of Peru for the purpose of collecting guano samples. The authors also thank the park rangers of the islands, responsible for the care of seabird colonies and the monitoring of guano accumulation, within the Subunit for the Extraction and Commercialization of Fertilizers of the Productive Rural Agricultural Development Program (AGRORURAL).

Data Availability

Data availability:The datasets generated and/or analyzed during the current study are publicly available in the NCBI BioProject repository under accession number PRJNA1454567. All other relevant data generated or analyzed during this study are included in this published article and its Supplementary Information files.

References

1. Cushman, G. T. *Studies in Environment and History: Guano and the Opening of the Pacific World: A Global Ecological History: A Global Ecological History* (Cambridge University Press, 2013). 10.1017/cbo9781139047470
2. O'Phelan, S. Humboldt, el Perú y sus recursos naturales: entre la plata y el guano. *Humboldt im Netz* 11, 73–82 (2010).
3. Anderson, W. B. & Polis, G. A. Nutrient fluxes from water to land: seabirds affect plant nutrient status on Gulf of California islands. *Oecologia* **118**, 324–332 (1999).
4. Weimerskirch, H., Bertrand, S., Silva, J., Bost, C. & Peraltilla, S. Foraging in Guanay cormorant and Peruvian booby, the major guano-producing seabirds in the Humboldt Current System. *Mar. Ecol. Prog Ser.* **458**, 231–245 (2012).
5. Sifuentes-García, Á., Zavalaga, C. B. & Lozano-Sanllehi, S. Estimación de la cantidad de guano de aves marinas acumulado en islas y puntas de Perú mediante un método geoestadístico en SIG. *Ecol. Austral*. **30**, 472–483 (2020).
6. Parco-Quispe, M., Camacho-Villanueva, A. A. & Parco-Quinchori, J. A. Dionisio-Saldaña, F. E. Efecto de niveles de aplicación de guano de islas en incremento de frutos de cacao. *Rev. Tecnol Marcha*. **35**, 105–114 (2022).
7. Godfray, H. C. J. et al. Food security: the challenge of feeding 9 billion people. *Science* **327**, 812–818 (2010).
8. Tilman, D., Balzer, C., Hill, J. & Befort, B. L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20260–20264 (2011).

9. Cosme, R. C., Reynoso, A. F. & Sanabria, S. Effect of island guano on the yield of two varieties of Quinoa (Chenopodium quinoa Willd) on degraded soil. *Agroindustrial Sci.* **10**, 191–198 (2020).
10. Olivera-Tacora, L. C., Salcedo-Mayta, S. M., Canihua-Rojas, J. & Quispe-Huincho, M. R. Cosme-De La Cruz, R. C. The use of cover crops with Vicia sativaand application of island guanoin the improving soil fertilityand yield of Chenopodium pallidicaule. *Agroindustrial Sci.* **12**, 141–146 (2022).
11. Huamán Villegas, F. D. Dosis óptima de guano de isla en el comportamiento agronómico del cultivo Brassica oleracea L.Var. itálica. *Rev. Cient. UNTRM Cienc. Nat. Ing.* **5**, 42–48 (2022).
12. Vásquez- Cantalicio, N. W. & Jacobo - Salinas, S. S. El guano de isla y su efecto en el rendimiento de la col (Brassica oleracea L) variedad lombarda (Capitata f. rubra) en Colicocha Huánuco. *ReInA* 2, 33–38 (2020).
13. Grzesiak, J. et al. A smelly business: Microbiology of Adélie penguin guano (Point Thomas rookery, Antarctica). *Sci. Total Environ.* **714**, 136714 (2020).
14. Mühldorfer, K. Bats and bacterial pathogens: a review: Bats and bacterial pathogens. *Zoonoses Public. Health*. **60**, 93–103 (2013).
15. Banskar, S., Bhute, S. S., Suryavanshi, M. V., Punekar, S. & Shouche, Y. S. Microbiome analysis reveals the abundance of bacterial pathogens in Rousettus leschenaultii guano. *Sci. Rep.* **6**, 36948 (2016).
16. Dimkić, I. et al. The microbiome of bat guano: for what is this knowledge important? *Appl. Microbiol. Biotechnol.* **105**, 1407–1419 (2021).
17. Otero, X. L., De La Peña-Lastra, S., Pérez-Alberti, A., Ferreira, T. O. & Huerta-Diaz, M. A. Seabird colonies as important global drivers in the nitrogen and phosphorus cycles. *Nat. Commun.* **9**, 246 (2018).
18. Newman, M. M., Kloepper, L. N., Duncan, M., McInroy, J. A. & Kloepper, J. W. Variation in bat guano bacterial community composition with depth. *Front. Microbiol.* **9**, 914 (2018).
19. Ellis, J. C. Marine birds on land: A review of plant biomass, species richness, and community composition in seabird colonies. *Plant. Ecol.* **181**, 227–241 (2005).
20. De Leon, M. P., Montecillo, A. D., Pinili, D. S., Siringan, M. A. T. & Park, D. S. Bacterial diversity of bat guano from Cabalyorisa Cave, Mabini, Pangasinan, Philippines: A first report on the metagenome of Philippine bat guano. *PLoS One*. **13**, e0200095 (2018).
21. Justel-Díez, M. et al. Inputs of seabird guano alter microbial growth, community composition and the phytoplankton-bacterial interactions in a coastal system. *Environ. Microbiol.* **25**, 1155–1173 (2023).
22. Keller, M. & Zengler, K. Tapping into microbial diversity. *Nat. Rev. Microbiol.* **2**, 141–150 (2004).
23. Riesenfeld, C. S., Schloss, P. D. & Handelsman, J. Metagenomics: genomic analysis of microbial communities. *Annu. Rev. Genet.* **38**, 525–552 (2004).
24. Klindworth, A. et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1–e1 (2013).

25. Schoch, C. L. et al. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc. Natl. Acad. Sci. U. S. A.* 109, 6241–6246 (2012).
26. Allison, S. D. & Martiny, J. B. H. Colloquium paper: resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* 105 Suppl 1, 11512–11519 (2008).
27. Louca, S., Parfrey, L. W. & Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**, 1272–1277 (2016).
28. Francioli, D., Lentendu, G., Lewin, S. & Kolb DNA metabarcoding for the characterization of terrestrial Microbiota-pitfalls and solutions. *Microorganisms* **9**, 361 (2021).
29. Fierer, N. & Jackson, R. B. The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* 103, 626–631 (2006).
30. Lauber, C. L., Hamady, M., Knight, R. & Fierer, N. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.* **75**, 5111–5120 (2009).
31. Oren, A. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Syst.* **4**, 2 (2008).
32. Kour, D. et al. Microbial biofertilizers: Bioresources and eco-friendly technologies for agricultural and environmental sustainability. *Biocatal. Agric. Biotechnol.* **23**, 101487 (2020).
33. Soto-López, J. D. et al. Taxonomic and functional profiling of bat guano microbiota from hiking trail-associated tunnels: a potential risk for human health? *Environ. Microbiome*. **20**, 123 (2025).
34. Zdanowski, M. K. et al. Bacterial diversity in AdÃ©lie penguin, Pygoscelis adeliae, guano: molecular and morpho-physiological approaches. *FEMS Microbiol. Ecol.* **50**, 163–173 (2004).
35. Mandic-Mulec, I., Stefanic, P. & van Elsas, J. D. Ecology of Bacillaceae. *Microbiol. Spectr.* **3**, TBS–0017 (2015).
36. Chen, Y. H., Shyu, Y. T. & Lin, S. S. Characterization of candidate genes involved in halotolerance using high-throughput omics in the halotolerant bacterium Virgibacillus chiguensis. *PLoS One*. **13**, e0201346 (2018).
37. Ahn, S. W., Lee, S. H., Son, H. S., Roh, S. W. & Choi, Y. E. Genomic analysis of halophilic bacterium, Lentibacillus sp. CBA3610, derived from human feces. *Gut Pathog*. **13**, 41 (2021).
38. Louca, S. et al. Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* **2**, 936–943 (2018).
39. Barbosa, A. et al. Age-Related Differences in the Gastrointestinal Microbiota of Chinstrap Penguins (Pygoscelis antarctica). *PLoS One*. **11**, e0153215 (2016).
40. Dewar, M. L. et al. Interspecific variations in the gastrointestinal microbiota in penguins. *Microbiologyopen* **2**, 195–204 (2013).
41. Bahram, M. et al. Structure and function of the global topsoil microbiome. *Nature* **560**, 233–237 (2018).

42. Tripathi, B. M. et al. Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *ISME J.* **12**, 1072–1083 (2018).
43. Lozupone, C. A. & Knight, R. Global patterns in bacterial diversity. *Proc. Natl. Acad. Sci. U. S. A.* 104, 11436–11440 (2007).
44. Rath, K. M., Fierer, N., Murphy, D. V. & Rousk, J. Linking bacterial community composition to soil salinity along environmental gradients. *ISME J.* **13**, 836–846 (2019).
45. Zhang, G. et al. Salinity controls soil microbial community structure and function in coastal estuarine wetlands. *Environ. Microbiol.* **23**, 1020–1037 (2021).
46. Egidi, E. et al. A few Ascomycota taxa dominate soil fungal communities worldwide. *Nat. Commun.* **10**, 2369 (2019).
47. Tedersoo, L. et al. Fungal biogeography. Global diversity and geography of soil fungi. *Science* **346**, 1256688 (2014).
48. Arné, P., Risco-Castillo, V., Jouvion, G., Le Barzic, C. & Guillot, J. Aspergillosis in wild birds. *J. Fungi (Basel)*. **7**, 241 (2021).
49. Wasti, I. G. et al. Fungal communities in bat guano, speleothem surfaces, and cavern water in Madai cave, Northern Borneo (Malaysia). *Mycology* **12**, 188–202 (2021).
50. Bohacz, J. Biodegradation of feather waste keratin by a keratinolytic soil fungus of the genus Chrysosporium and statistical optimization of feather mass loss. *World J. Microbiol. Biotechnol.* **33**, 13 (2017).
51. Kitowski, I., Korniłłowicz-Kowalska, T., Bohacz, J. & Ciesielska, A. Dispersal of Aphanoascus keratinophilus by the rook Corvus frugilegus during breeding in East Poland. *Sci. Rep.* **12**, 2142 (2022).
52. Ozimek, E. & Hanaka, A. Mortierella species as the plant growth-promoting fungi present in the agricultural soils. *Agriculture* **11**, 7 (2020).
53. Telagathoti, A., Probst, M. & Peintner, U. Habitat, snow-cover and soil pH, affect the distribution and diversity of Mortierellaceae species and their associations to bacteria. *Front. Microbiol.* **12**, 669784 (2021).
54. Richardson, M. J. The coprophilous succession. in Fungal Succession (eds Hyde, K. D. & Jones, E. B. G.) 101–111 (Fungal Diversity, (2002).
55. Richardson, M. J. Diversity and occurrence of coprophilous fungi. *Mycol. Res.* **105**, 387–402 (2001).
56. Louca, S. et al. High taxonomic variability despite stable functional structure across microbial communities. *Nat. Ecol. Evol.* **1**, 15 (2016).
57. Nguyen, N. H. et al. FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol.* **20**, 241–248 (2016).
58. Wasimuddin et al. Evaluation of primer pairs for microbiome profiling from soils to humans within the One Health framework. *Mol. Ecol. Resour.* **20**, 1558–1571 (2020).

59. White, T. J. Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics. in PCR Protocols, a Guide to Methods and Applications (Scientific Research Publishing, (1990).
60. De Op, M. et al. Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. *PLoS One*. **9**, e97629 (2014).
61. Straub, D. et al. Interpretations of environmental microbial community studies are biased by the selected 16S rRNA (gene) amplicon sequencing pipeline. *Front. Microbiol.* **11**, 550420 (2020).
62. Di Tommaso, P. et al. Nextflow enables reproducible computational workflows. *Nat. Biotechnol.* **35**, 316–319 (2017).
63. Grüning, B. et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat. Methods*. **15**, 475–476 (2018).
64. da Leprevost, V. BioContainers: an open-source and community-driven framework for software standardization. *Bioinformatics* **33**, 2580–2582 (2017).
65. Andrews, S. & Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. (2010). https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
66. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
67. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods*. **13**, 581–583 (2016).
68. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
69. Nilsson, R. H. et al. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* **47**, D259–D264 (2019).
70. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
71. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. **8**, e61217 (2013).
72. Wickham, H. ggplot2: Elegant graphics for data analysis. *Springer-Verlag* (2016). https://ggplot2.tidyverse.org/
73. Kassambara, A. *ggpubr: ggplot2 based publication ready plots* (R package version 0.6.0). (2023). https://CRAN.R-project.org/package=ggpubr
74. Slowikowski, K. *ggrepel: Automatically position non-overlapping text labels with ggplot2* (R package version 0.9.4). *Comprehensive R Archive Network (CRAN)* (2023). https://CRAN.R-project.org/package=ggrepel
75. Neuwirth, E. *RColorBrewer: ColorBrewer palettes* (R package version 1.1-3). *Comprehensive R Archive Network (CRAN)* (2022). https://CRAN.R-project.org/package=RColorBrewer
76. Larsson, J. *eulerr: Area-proportional Euler and Venn diagrams with ellipses* (R package version 7.0.0). *Comprehensive R Archive Network (CRAN)* (2021). https://CRAN.R-project.org/package=eulerr

77. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
78. Wei, T. & Simko, V. *corrplot: Visualization of a correlation matrix* (R package version 0.95). *Comprehensive R Archive Network (CRAN)* (2021). https://CRAN.R-project.org/package=corrplot
79. Liu, C. microeco: Microbial Community Ecology Data Analysis. *Comprehensive R Archive Network (CRAN)* (2026). https://CRAN.R-project.org/package=microeco
80. Kolde, R. *pheatmap: Pretty heatmaps* (R package version 1.0.12). *Comprehensive R Archive Network (CRAN)* (2019). https://CRAN.R-project.org/package=pheatmap
81. Pedersen, T. L. *patchwork: The composer of plots* (R package version 1.2.0). *Comprehensive R Archive Network (CRAN)* (2022). https://CRAN.R-project.org/package=patchwork

Figures

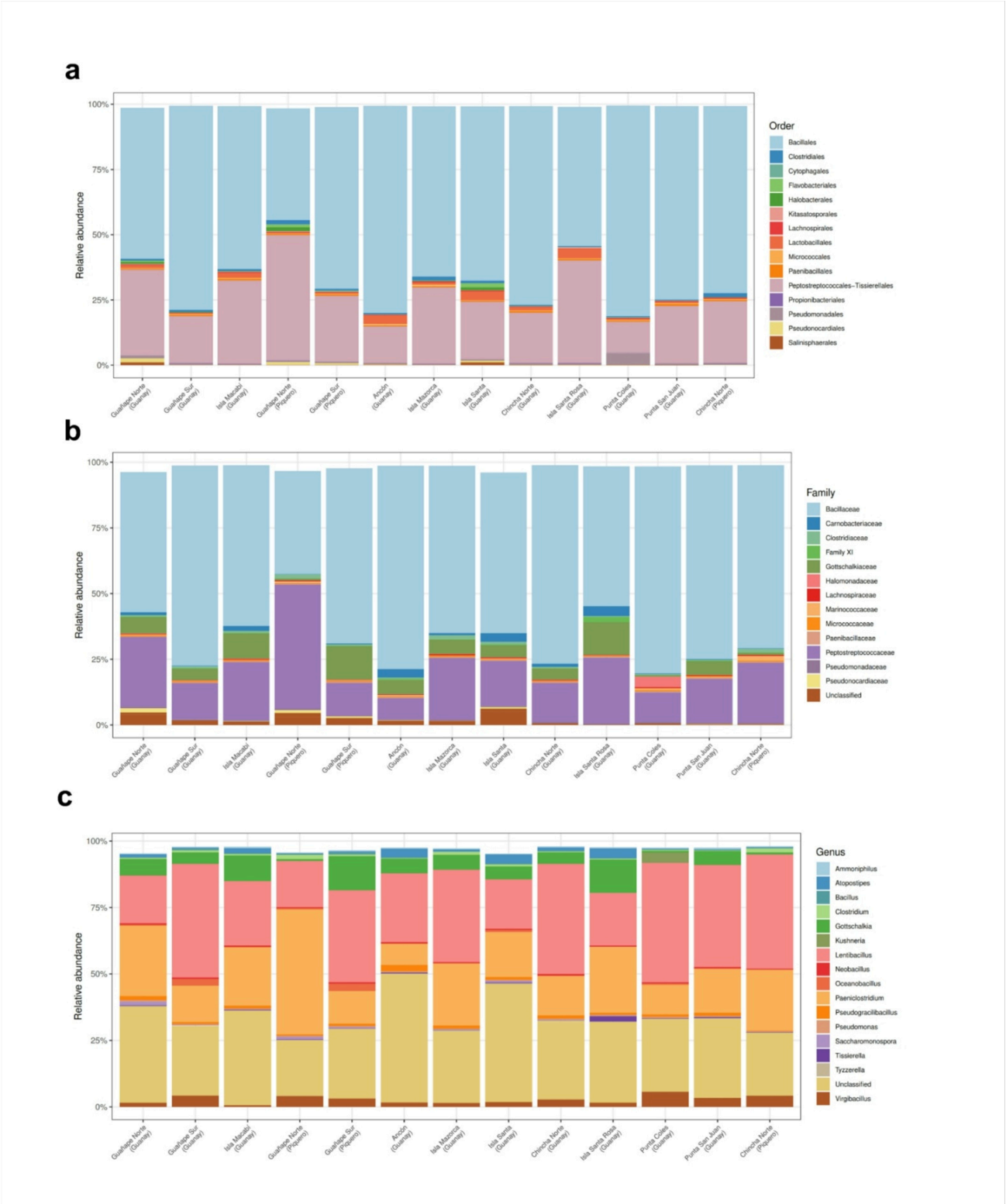


Figure 1

Relative taxonomic composition of bacterial communities associated with island guano, determined by 16S rRNA marker sequencing. (A) Order-level composition by island or guano headland. (B) Family-level composition by island or guano headland. (C) Genus-level composition by island or guano headland. Each stacked bar represents a composite sample corresponding to a sampling locality and the bird species associated with the guano, expressing the relative abundance of the principal bacterial taxa

identified. Colors indicate the different taxonomic groups at each classification level. The figure illustrates the dominance of certain bacterial lineages and the variation in their relative abundance across guano localities.

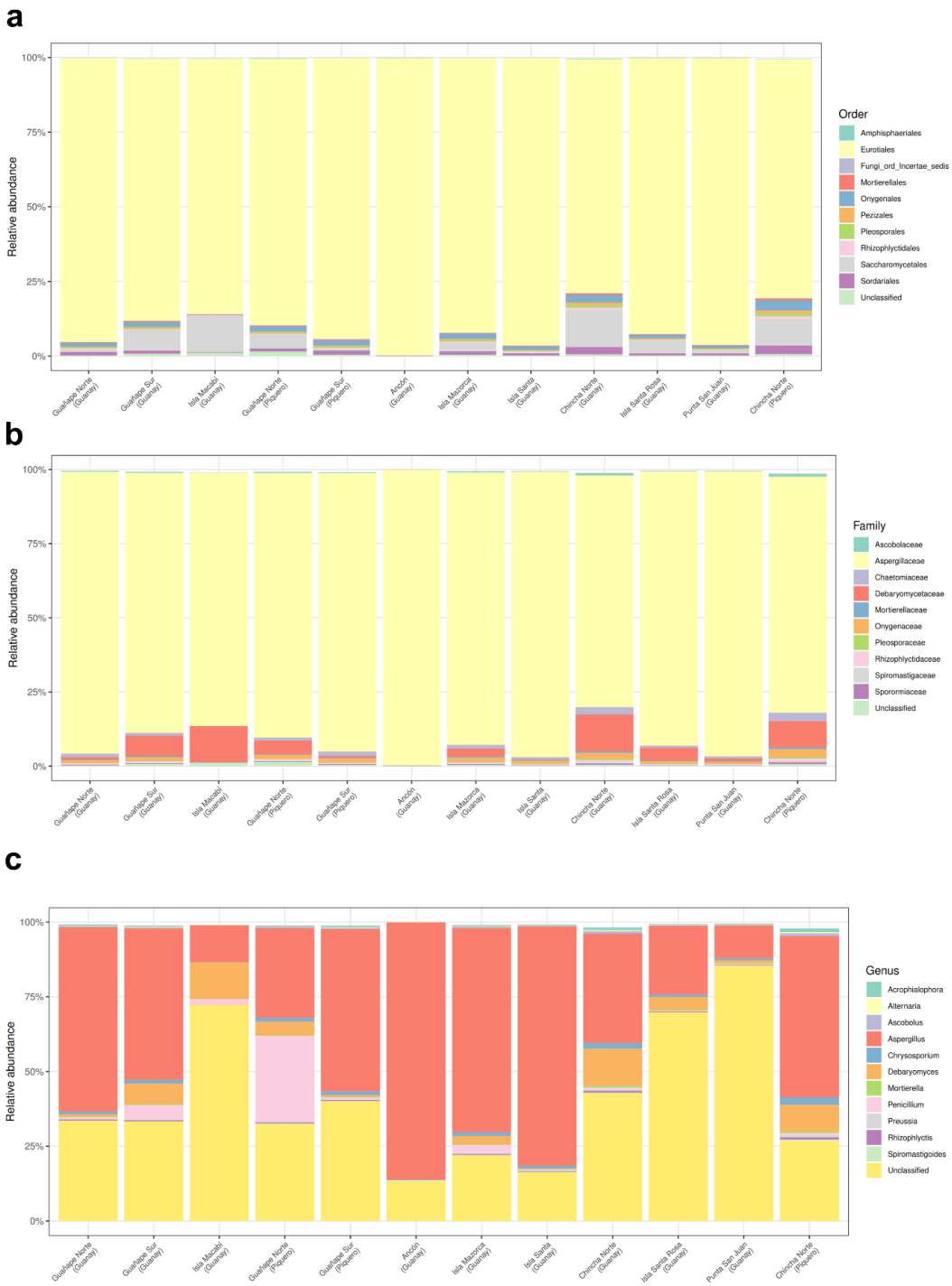


Figure 2

Relative taxonomic composition of fungal communities associated with island guano, determined by ITS marker sequencing. (A) Order-level composition by island or guano headland. (B) Family-level

composition by island or guano headland. (C) Genus-level composition by island or guano headland. Each stacked bar represents a composite sample corresponding to a sampling locality and the bird species associated with the guano, expressing the relative abundance of the principal fungal taxa identified. Colors indicate the different taxonomic groups at each classification level. The figure illustrates the dominance of specific fungal lineages and the variation in their relative abundance across guano localities.

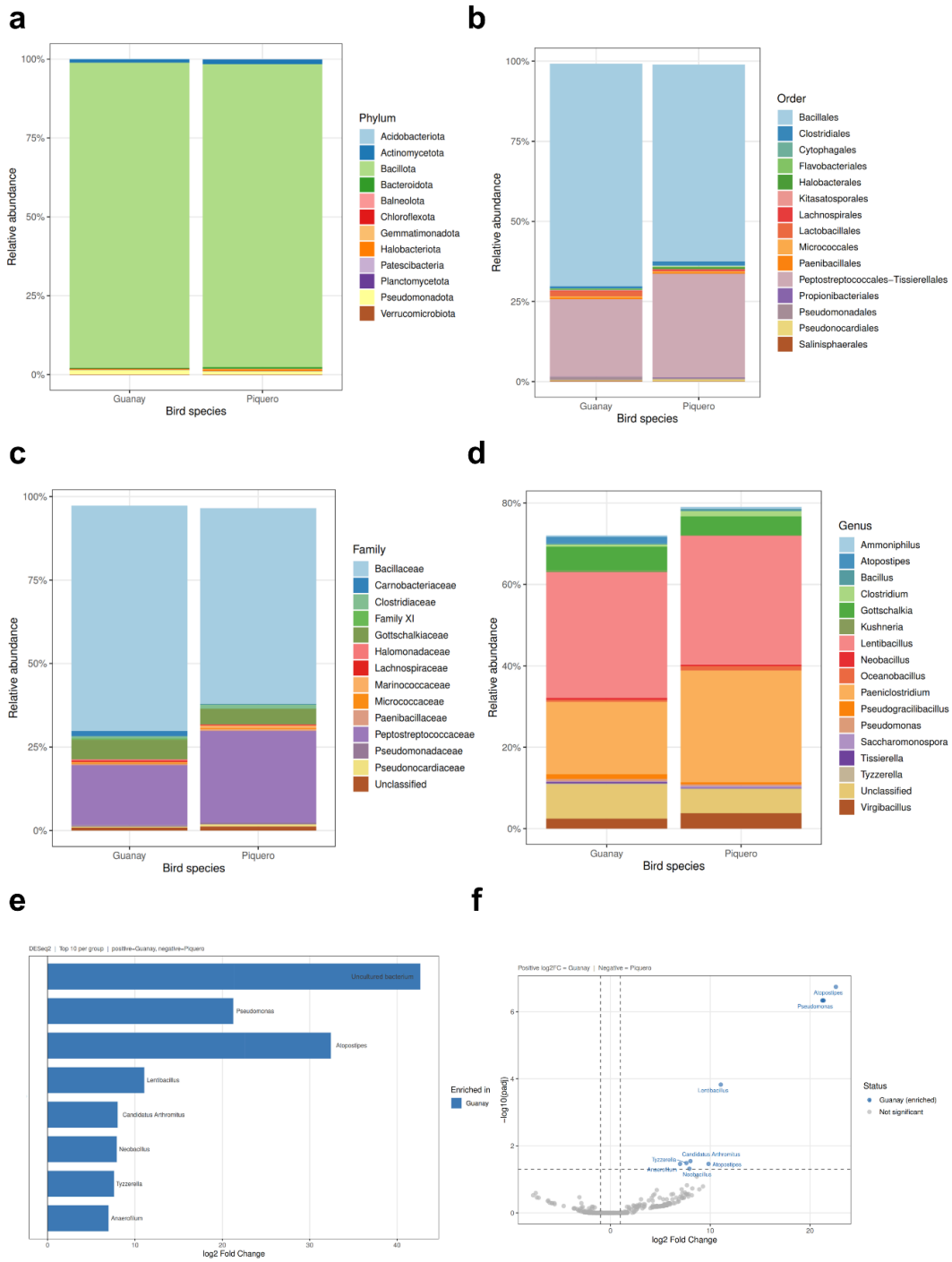


Figure 3

Taxonomic composition and differential abundance of bacterial communities associated with Guanay cormorant and Peruvian booby guano, determined by 16S rRNA marker sequencing. (A) Relative taxonomic composition at the phylum level. (B) Relative taxonomic composition at the order level. (C) Relative taxonomic composition at the family level. (D) Relative taxonomic composition at the genus level. (E) Bacterial taxa with the greatest differential abundance between Guanay cormorant and Peruvian booby guano, expressed as $\log_2$ fold-change. (F) Volcano plot of the differential abundance analysis, where each point represents a bacterial taxon; positive $\log_2$fold-change values indicate enrichment in Guanay cormorant and negative values indicate enrichment in Peruvian booby. Colored points correspond to taxa with significant differences, while grey points represent non-significant taxa.

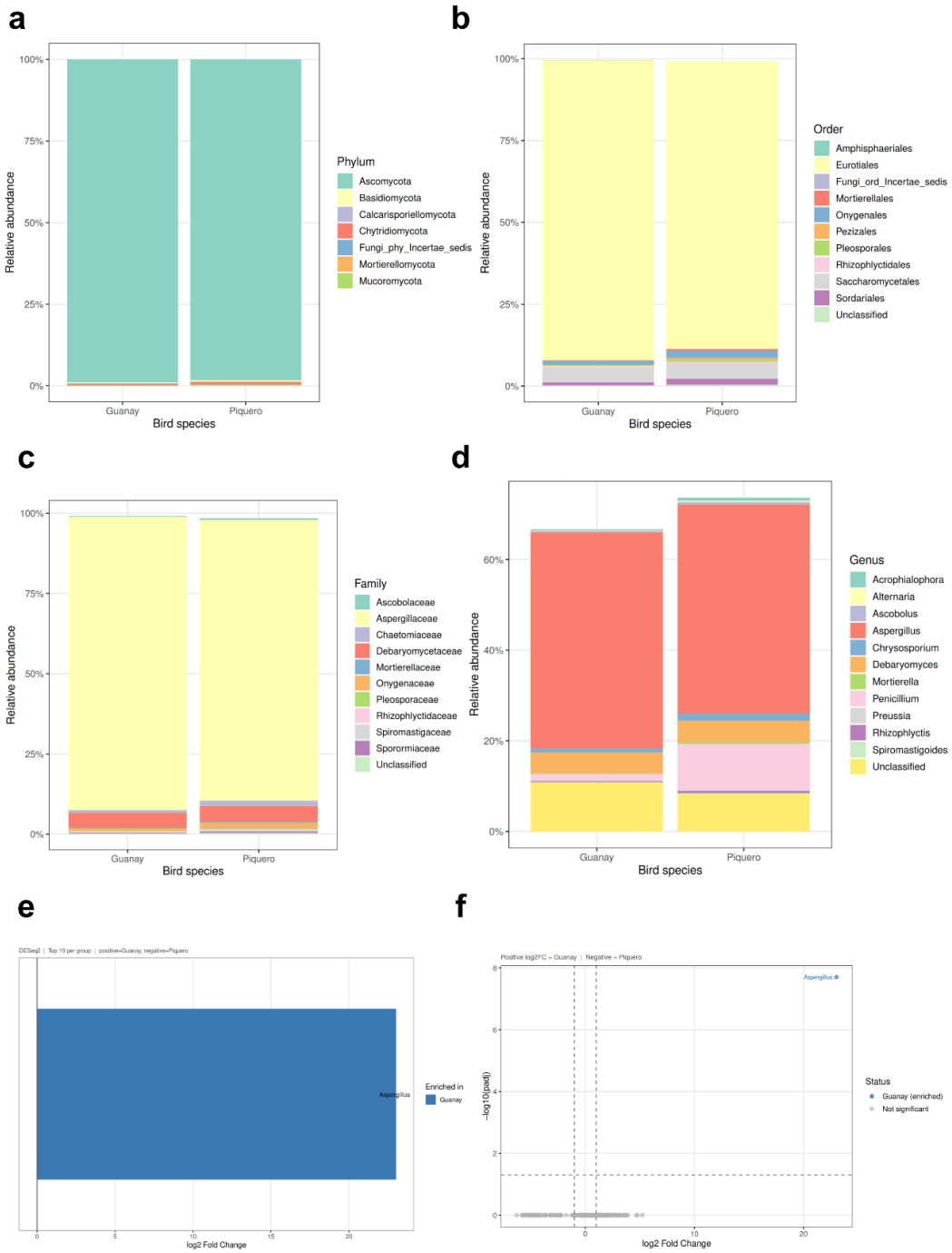


Figure 4

Taxonomic composition and differential abundance of fungal communities associated with Guanay cormorant and Peruvian booby guano, determined by ITS marker sequencing. (A) Relative taxonomic composition at the phylum level. (B) Relative taxonomic composition at the order level. (C) Relative taxonomic composition at the family level. (D) Relative taxonomic composition at the genus level. (E) Fungal taxa with the greatest differential abundance between Guanay cormorant and Peruvian booby

guano, expressed as log_2 fold-change. (F) Volcano plot of the differential abundance analysis, where each point represents a fungal taxon; positive log_2fold-change values indicate enrichment in Guanay cormorant and negative values indicate enrichment in Peruvian booby. Colored points correspond to taxa with significant differences, while grey points represent non-significant taxa.

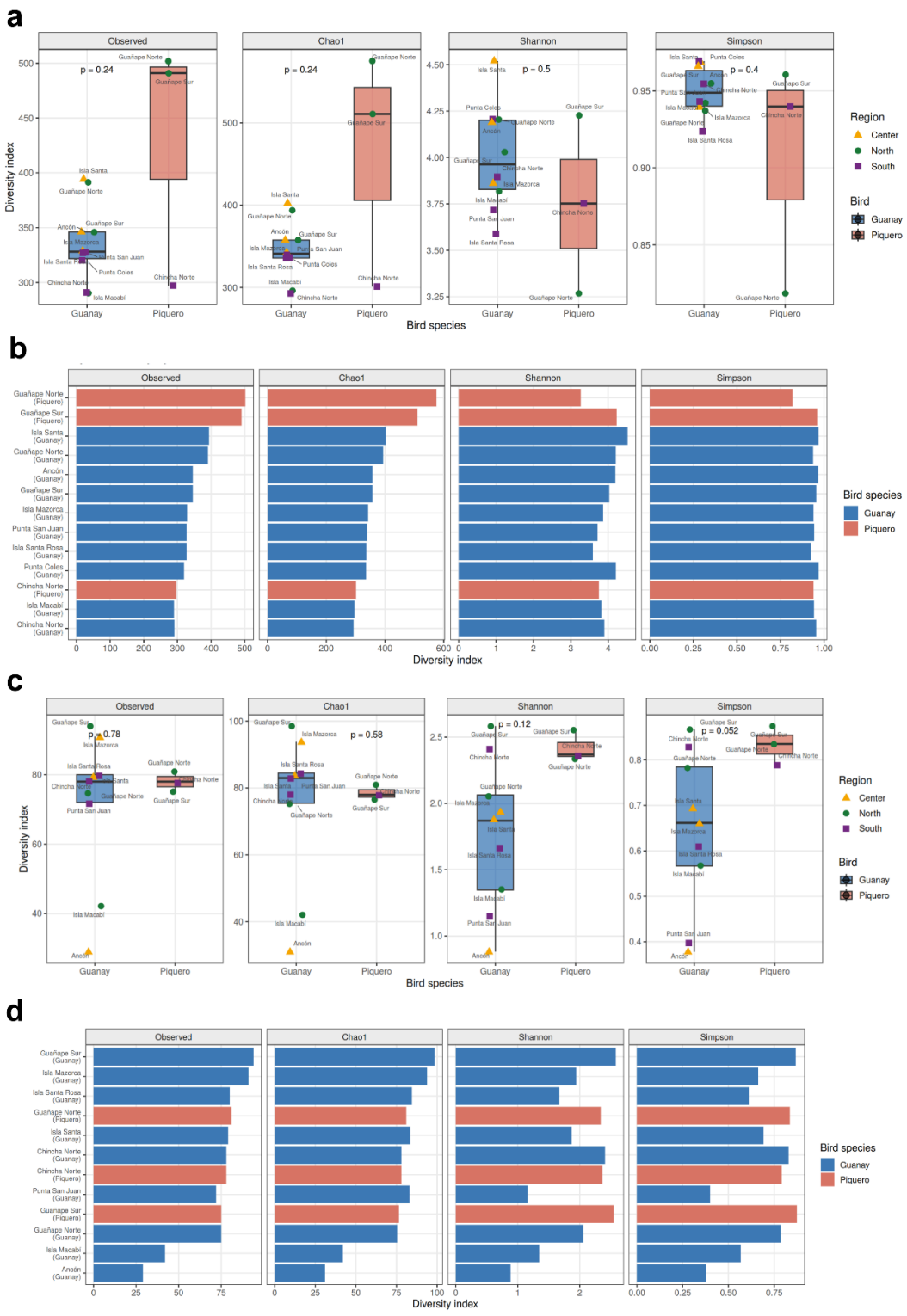


Figure 5

Alpha diversity analysis of bacterial and fungal communities associated with island guano. **(A)** Bacterial (16S) alpha diversity compared between Guanay cormorant and Peruvian booby guano using the Observed, Chao1, Shannon, and Simpson indices. **(B)** Alpha diversity values by island for bacteria (16S) using the Observed, Chao1, Shannon, and Simpson indices. **(C)** Fungal (ITS) alpha diversity compared between Guanay cormorant and Peruvian booby guano using the Observed, Chao1, Shannon, and Simpson indices. **(D)** Alpha diversity values by island for fungi (ITS) using the Observed, Chao1, Shannon, and Simpson indices. Each point represents a composite sample, and colors/symbols indicate the geographic region and bird species associated with the guano. P-values correspond to between-group comparisons; n.s., not significant.

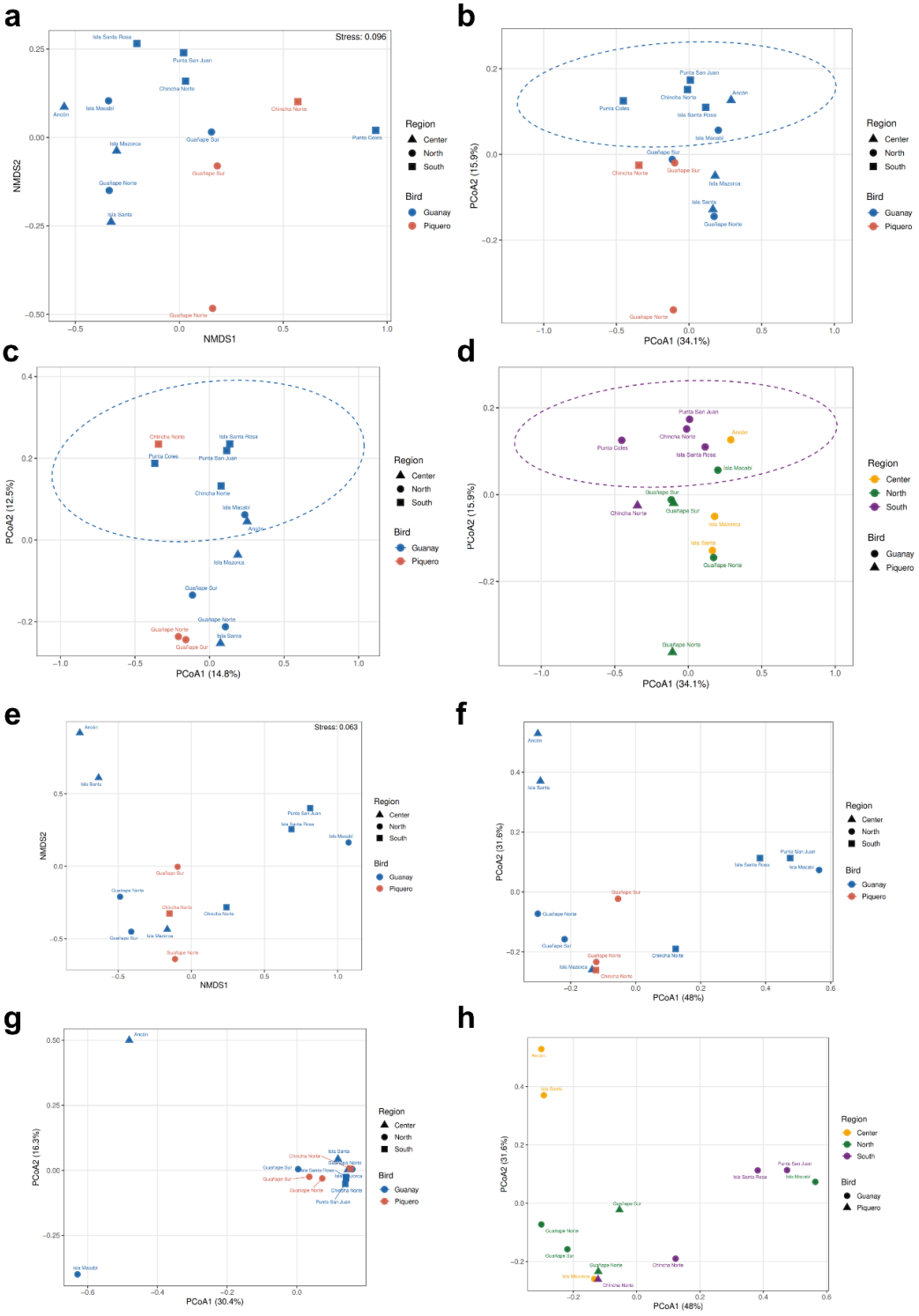


Figure 6

Beta diversity analysis of bacterial and fungal communities associated with island guano using dissimilarity metrics and multivariate ordination. (A) Bray–Curtis-based NMDS for bacteria (16S). (B) Bray–Curtis-based PCoA for bacteria (16S). (C) Jaccard-based PCoA for bacteria (16S). (D) PCoA of bacterial communities (16S) showing sample distribution by geographic region. (E) Bray–Curtis-based NMDS for fungi (ITS). (F) Bray–Curtis-based PCoA for fungi (ITS). (G) Jaccard-based PCoA for fungi

(ITS). (H) PCoA of fungal communities (ITS) showing sample distribution by geographic region. Each point represents a composite guano sample; colors and symbols indicate the bird species associated with the guano and/or the geographic region of origin, as shown in the legend of each panel. Percentages indicated on PCoA axes correspond to the variance explained by each principal coordinate, while stress values in NMDS analyses indicate the ordination fit. Dashed ellipses represent group dispersion in multivariate space.

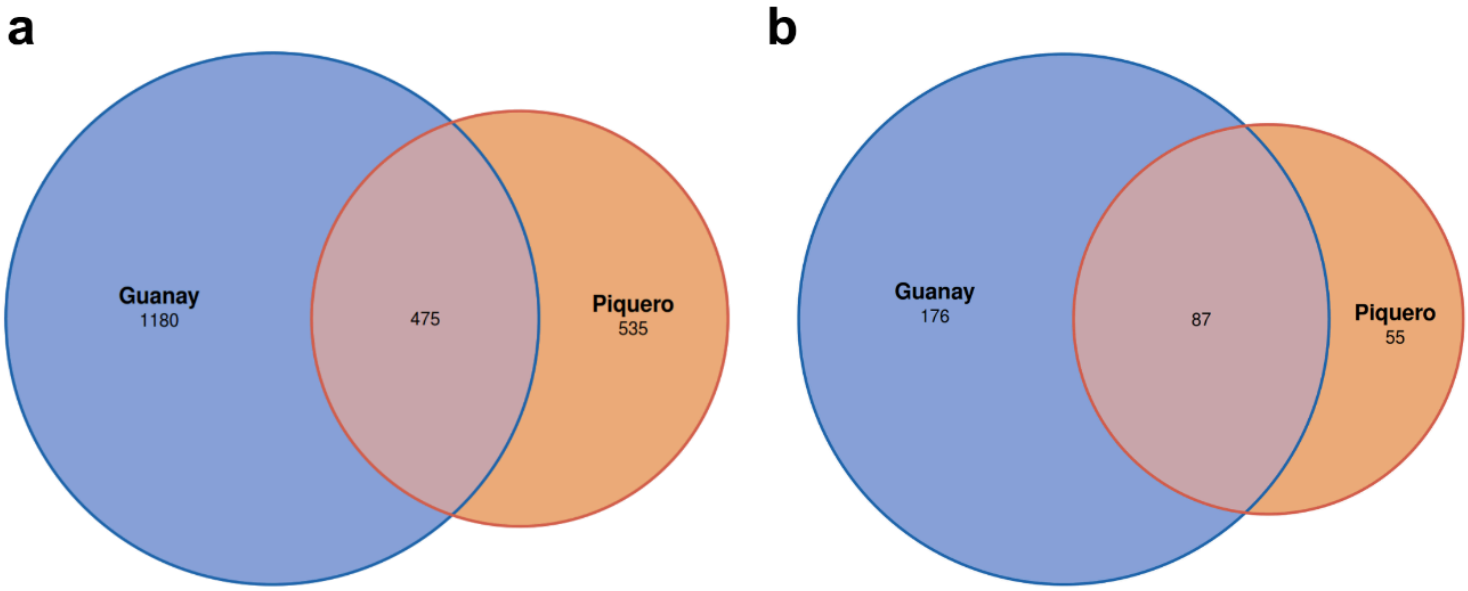


Figure 7

Distribution of shared and exclusive ASVs between Guanay cormorant and Peruvian booby guano. **(A)** Bacterial amplicon sequence variants (ASVs) detected using the 16S rRNA marker. **(B)** Fungal ASVs detected using the ITS marker. Venn diagrams show the number of ASVs exclusive to each bird species and the number of ASVs shared between both guano types, allowing visualization of the common microbial core and the taxonomically exclusive fraction of each group.

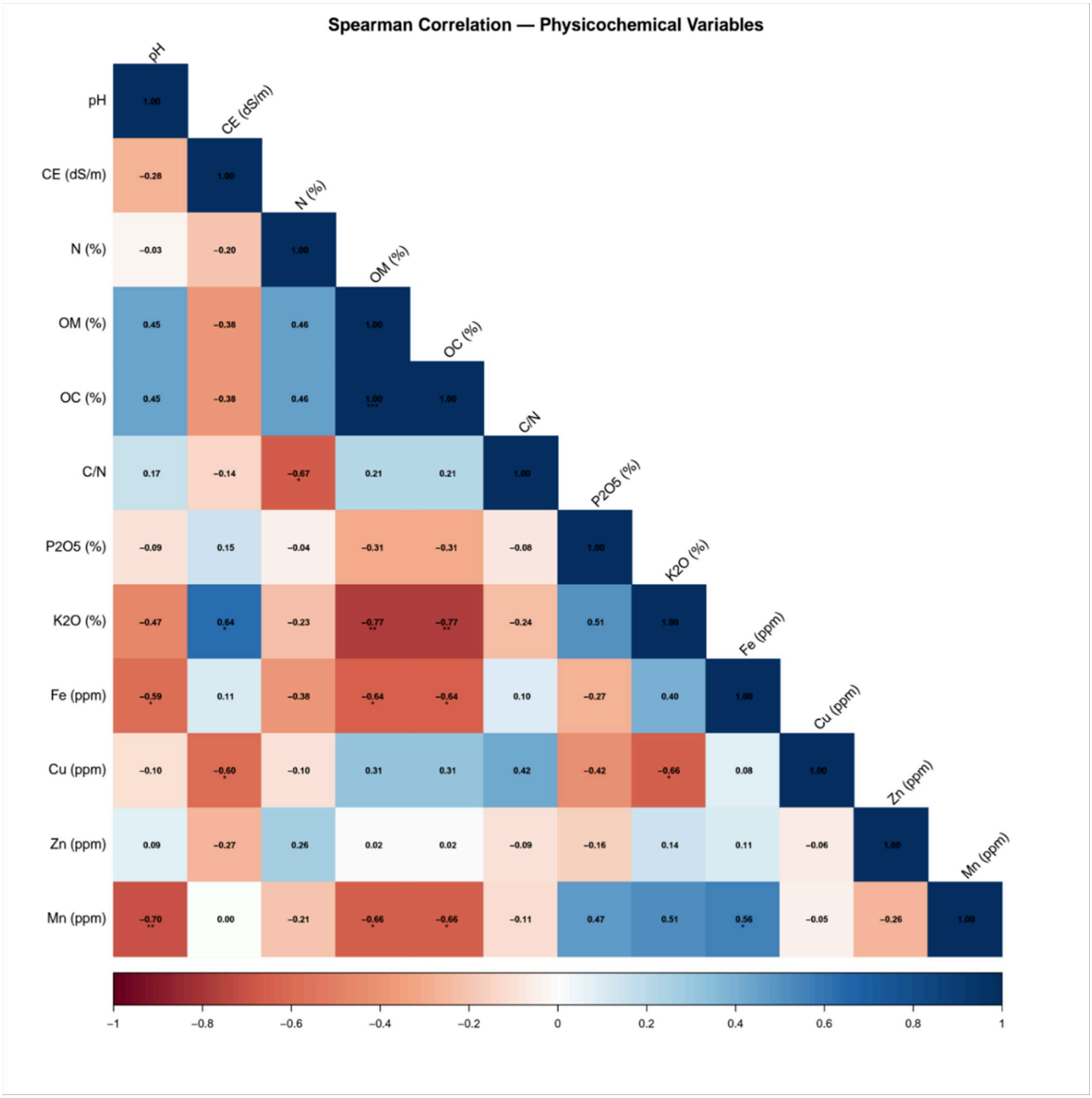


Figure 8

Spearman correlation matrix among physicochemical variables of island guano. Each cell shows the Spearman correlation coefficient (ρ) between pairs of variables, including pH, electrical conductivity (EC), total nitrogen (N), organic matter (OM), organic carbon (OC), C/N ratio, phosphorus (P_2O_5), potassium (K_2O), iron (Fe), copper (Cu), zinc (Zn), and manganese (Mn). Color intensity indicates the magnitude and direction of the correlation, ranging from negative values (red) to positive values (blue), as shown in the

scale. Numerical values within each cell correspond to the correlation coefficient, and asterisks indicate statistically significant associations between variables.

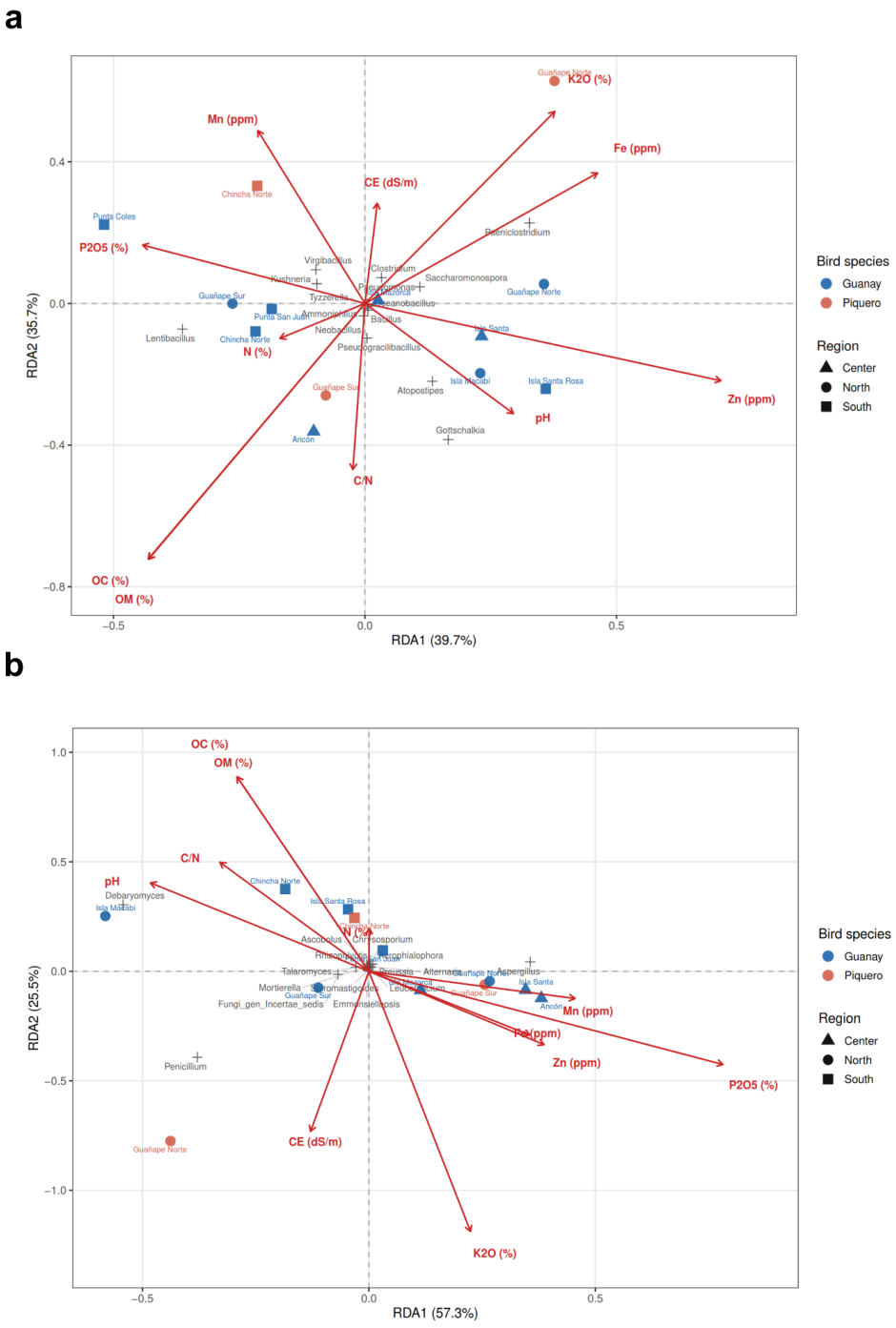


Figure 9

Redundancy analysis (RDA) of island guano microbial communities in relation to physicochemical variables. **(A)** RDA for the 15 most abundant bacterial genera from the 16S dataset. **(B)** RDA for the 15 most abundant fungal genera from the ITS dataset. Points represent composite guano samples, colored

by associated bird species (Guanay cormorant and Peruvian booby) and differentiated by symbol according to geographic region (central, northern, and southern). Black crosses indicate the position of dominant genera in ordination space. Red arrows represent the physicochemical variables evaluated; their direction and length reflect the relative contribution of each variable to the ordination pattern. Percentages on the axes correspond to the variance explained by the first two canonical axes. The figure allows visualization of the association between microbial composition and physicochemical gradients in guano.

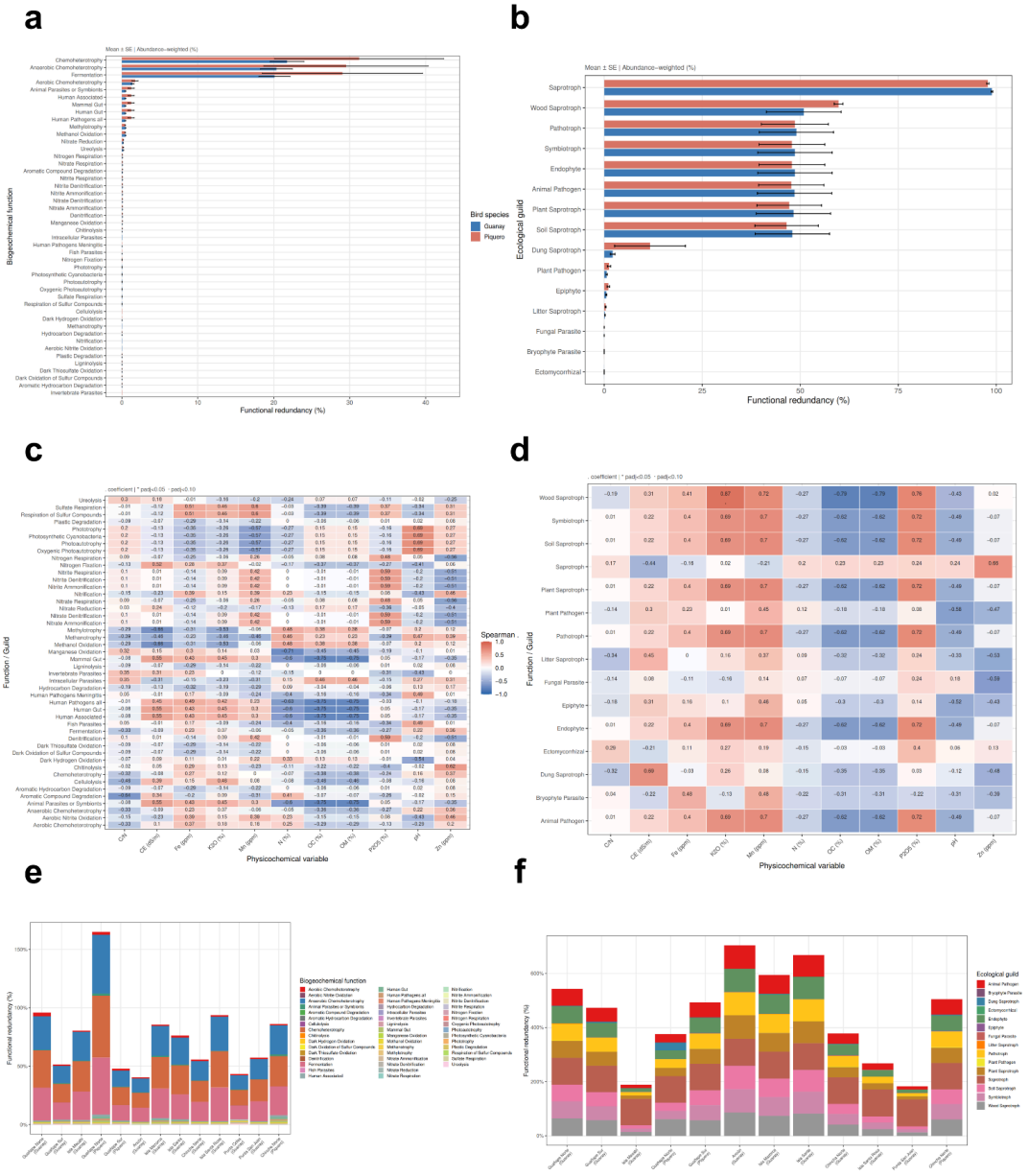


Figure 10

Predicted functional redundancy of bacterial and fungal communities associated with island guano, and their relationship with physicochemical variables. **(A)** Functional redundancy of bacteria (16S) according to biogeochemical categories predicted by FAPROTAX, compared between Guanay cormorant and Peruvian booby guano. **(B)** Functional redundancy of fungi (ITS) according to ecological guilds predicted by FUNGuild, compared between Guanay cormorant and Peruvian booby guano. In both panels, bars represent the mean ± standard error of the weighted abundance percentage. **(C)**Heatmap of correlations between bacterial biogeochemical functions predicted by FAPROTAX and guano physicochemical variables. **(D)** Heatmap of correlations between fungal ecological guilds predicted by FUNGuild and guano physicochemical variables. In the heatmaps, colors indicate the direction and magnitude of the correlation, and numerical values within each cell correspond to correlation coefficients. **(E)** Distribution of bacterial biogeochemical functions predicted by FAPROTAX across each island or guano headland. **(F)** Distribution of fungal ecological guilds predicted by FUNGuild across each island or guano headland. In panels E and F, stacked bars represent the estimated functional redundancy percentage for each composite guano sample.

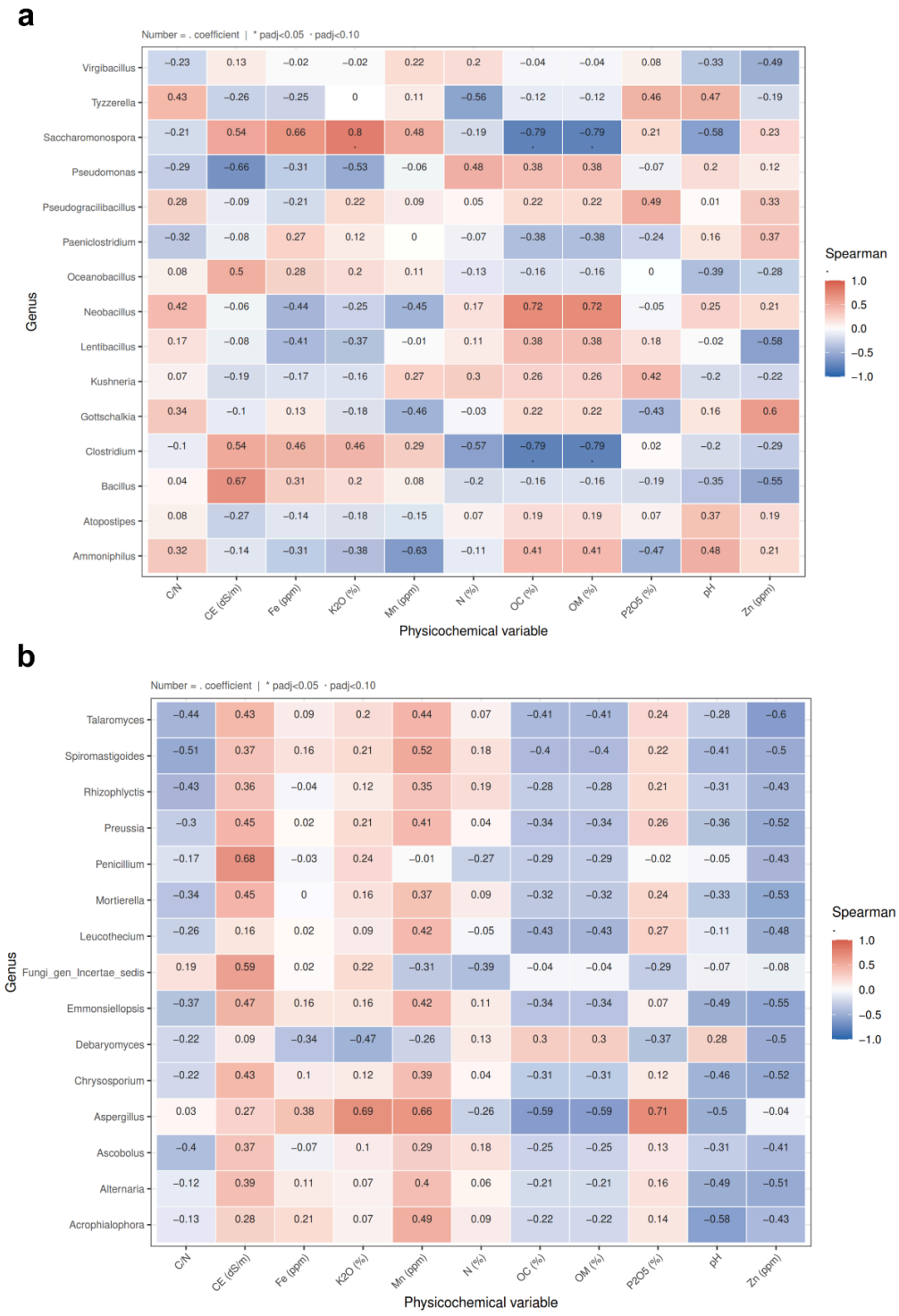


Figure 11

Spearman correlations between dominant microbial genera and physicochemical variables of island guano. (A) Spearman correlations for the principal bacterial genera from the 16S dataset. (B) Spearman correlations for the principal fungal genera from the ITS dataset. In both panels, rows represent microbial genera and columns represent the physicochemical variables evaluated (C/N, EC, Fe, K_2O, Mn, N, OC, OM, P_2O_5, pH, and Zn). Numerical values within each cell correspond to the Spearman correlation

coefficient (ρ), while the color scale indicates the direction and magnitude of the association, ranging from negative correlations (blue) to positive correlations (red). Symbols indicate the adjusted statistical significance level: * p_adj < 0.05 and • p_adj < 0.10.

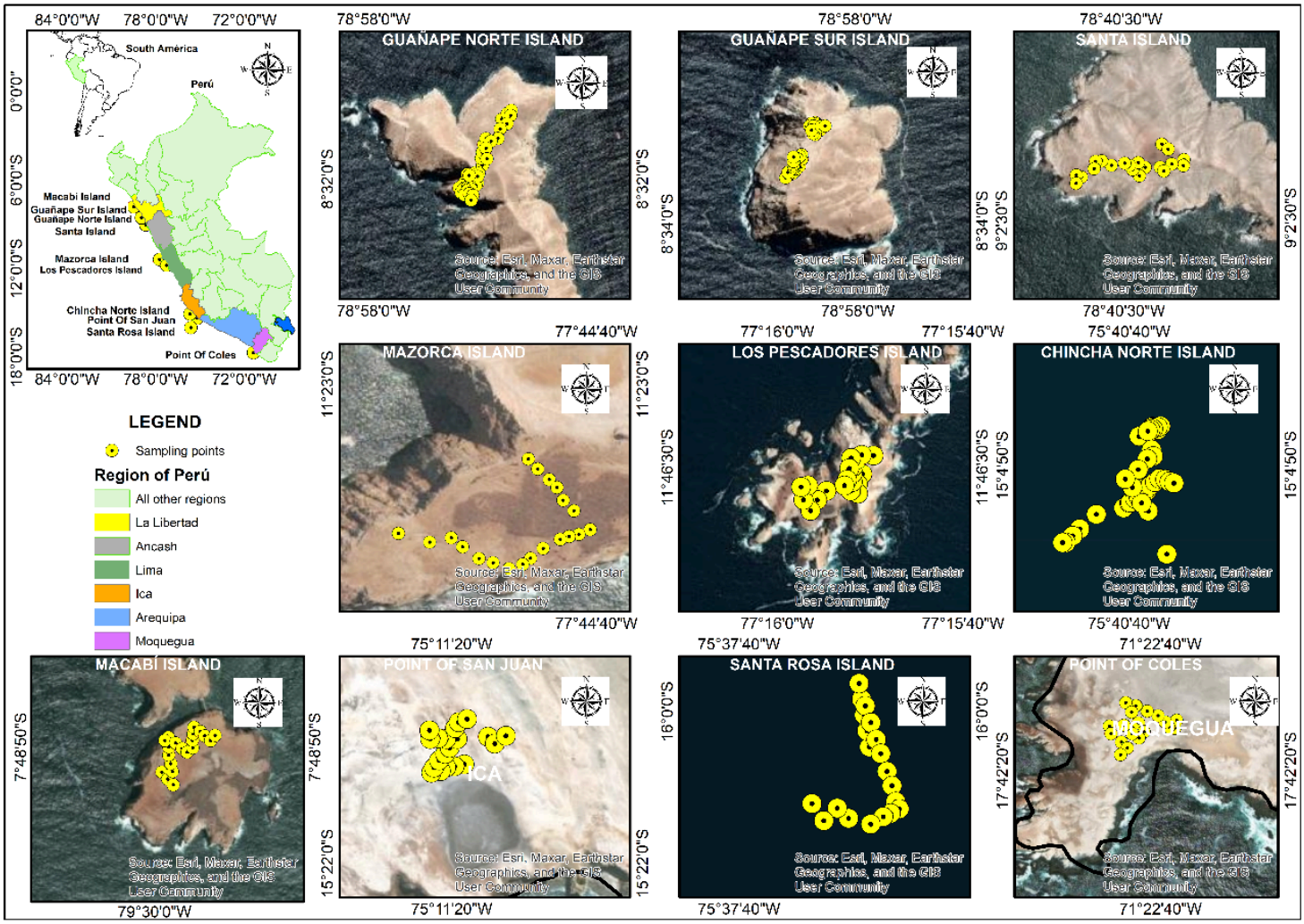


Figure 12

Map of marine bird guano sampling sites across eight (8) islands and two (2) guano headlands along the Peruvian coastline. At each island and headland, 20 individual samples were collected for each dominant bird species, yielding a total of 260 sampling points.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuppMaterialMetagenomicaenGuanodeislaINIAPer2026.docx