Acta Tropica 280 (2026) 108209

Contents lists available at ScienceDirect

Acta Tropica

journal homepage: www.elsevier.com/locate/actatropica

Ecological divergence of *Paracoccidioides brasiliensis* lineages and *P. lutzii* with SDG-relevant implications

Attila J. Trájer *, Katalin Vehovszky

Sustainability Solutions Research Lab, University of Pannonia, H–8200 Veszprém, Egyetem utca 10, Hungary

ARTICLE INFO

Keywords:
Paracoccidioidomycosis
Ecological niche modelling
Bioclimatic variables
Environmental clustering
South America
Machine learning

ABSTRACT

The ecological differentiation of *Paracoccidioides brasiliensis* lineages (S1, PS3, S1a/b) and *Paracoccidioides lutzii* across South America was analysed using environmental, climatic, edaphic, and anthropogenic data. Occurrences were linked to elevation, urbanisation, Köppen–Geiger climate classes, and bioclimatic variables. S1 exhibited the broadest ecological range, spanning mid-elevation uplands to lowland Amazonia and multiple climate zones (Af, Am, Aw, Cfa). PS2, PS3 and *P. lutzii* were mainly restricted to mid-elevation environments. PS2 was associated with slightly higher elevations; PS3 overlapped broadly with S1, and *P. lutzii* showed a stronger association with ustult ultisols and tropical savannah (Aw) climates. Machine learning models separated *P. brasiliensis* from *P. lutzii* with high accuracy, whereas lineage-level discrimination was more variable, particularly for PS2 versus *P. lutzii*. Temperature variables (bio11, bio6, bio1), elevation, and soil type were the strongest predictors, with precipitation and aridity playing secondary roles. Ordination and niche analyses revealed strong overlap among lineages. S1 and PS3 lineages showed the greatest overlap; PS2 was more distinct, and *P. lutzii* occupied an intermediate position. High niche similarity with the xenarthran host *Dasypus novemcinctus* was also observed. Overall, ecological differentiation within *Paracoccidioides* reflects gradient-driven structuring rather than discrete niches, with implications for spatial risk assessment, particularly for S1-associated regions, and targeted environmental management aligned with SDGs 3, 13, and 15.

1. Introduction

Fungal diseases are increasingly recognised as major global health challenges, causing substantial morbidity and mortality worldwide (Benedict et al., 2017; Denning, 2024). Climate change, land-use transformation, environmental disturbance, and increasing global mobility are further altering the emergence and spread of pathogenic fungi, emphasizing the importance of understanding their ecological drivers and environmental distribution patterns (Gottdenker et al., 2014; Rokas, 2022; Williams et al., 2024; Zhou et al., 2024).

Paracoccidioidomycosis (PCM) is one of the most important endemic systemic mycoses in Latin America and remains the leading cause of death from fungal infections among immunocompetent individuals in Brazil (Rodrigues and Nosanchuk, 2021). The etiological agents belong to the genus *Paracoccidioides*, thermally dimorphic soil-associated fungi within the order Onygenales (San-Blas et al., 2002; Hrycyk et al., 2018). Among these, *Paracoccidioides brasiliensis* (Splend.) Almeida (Almeida, 1930) and *Paracoccidioides lutzii* Vilela, de Hoog, Bagagli & L. Mend. are the principal taxa associated with human disease in South America (Teixeira et al., 2014). Their distribution is influenced by climatic conditions, soil properties, environmental disturbance, and mammalian reservoir hosts (Restrepo-Moreno, 2018).

Molecular and phylogenetic studies have demonstrated that *P. brasiliensis* represents a complex of genetically differentiated phylogenetic lineages rather than a single homogeneous species (Bagagli et al., 2008; Muñoz et al., 2016; Cocio et al., 2020). The major phylogenetic lineages include S1, PS2 (*Paracoccidioides americana*), PS3 (*Paracoccidioides restrepiensis*), and PS4 (*Paracoccidioides venezuelensis*) (Turissini et al., 2017; Pinheiro et al., 2020; Andrade-Silva et al., 2021). Among these, S1 is the most widespread and epidemiologically important lineage, whereas PS2–PS4 exhibit more restricted geographic distributions (Andrade-Silva et al., 2021). Previous studies further suggest differences among lineages in ecological tolerance, serological characteristics, and population structure (Sharpton et al., 2009; Hahn et al., 2022). Evidence of gene flow and hybridization within parts of the complex also indicates that environmental heterogeneity and geographic barriers may contribute to lineage diversification (Cocio et al., 2020; Teixeira et al., 2020).

* Corresponding author.
E-mail address: trajer.attila@mk.uni-pannon.hu (A.J. Trájer).

https://doi.org/10.1016/j.actatropica.2026.108209
Received 26 January 2026; Received in revised form 4 June 2026; Accepted 1 July 2026
Available online 2 July 2026

Paracoccidioides species are generally associated with humid, well-aerated soils in tropical and subtropical environments, including forests, riparian habitats, agricultural areas, and disturbed vegetation (Terçarioli et al., 2007; Fabris et al., 2024). Environmental factors such as soil moisture, precipitation seasonality, and temperature strongly influence fungal persistence, sporulation, and dispersal (Terçarioli et al., 2007).

Because South America encompasses pronounced climatic and topographic heterogeneity, environmental filtering likely plays a central role in determining the spatial distribution and ecological limits of individual lineages (Fabris et al., 2024; Calle et al., 2001). However, despite increasing phylogenetic knowledge, the ecological differentiation and environmental niche structure of *Paracoccidioides* lineages remain poorly understood.

The nine-banded armadillo, *Dasypus novemcinctus* Linnaeus, 1758, is considered the principal reservoir host and an important ecological indicator for *Paracoccidioides* spp. (Bagagli et al., 1998, 2003, 2021; Cocio et al., 2020). The fungus has frequently been detected in armadillo tissues and in soil collected from their burrows (Arantes et al., 2016; Hrycyk et al., 2018). Experimental evidence further suggests that these hosts function primarily as transient environmental reservoirs rather than sites of persistent infection, reinforcing the ecological connection between fungal occurrence and environmental conditions (Richini-Pereira et al., 2009).

Such gaps limit accurate prediction of disease risk and potential distribution shifts under ongoing climatic and land-use change (Rodrigues and Nosanchuk, 2021; Hrycyk et al., 2018). Integrating climatic, edaphic, host-association, and phylogenetic data within a multivariate ecological framework may therefore improve understanding of lineage-specific environmental tolerances and distribution patterns (Cocio et al., 2020; Arantes et al., 2016; Hahn et al., 2022).

The present study investigates the ecological differentiation of *Paracoccidioides* species and phylogenetic lineages across South America using environmental ordination, clustering analyses, machine learning approaches, climatic suitability modelling, and quantitative niche-overlap metrics. By integrating environmental, phylogenetic, and host-association data, the study aims to identify lineage-specific environmental patterns, evaluate major ecological drivers, and clarify the broader ecological and epidemiological implications of diversification within the *Paracoccidioides* complex.

2. Materials and methods

2.1. Study workflow

To investigate how environmental, climatic, edaphic, and anthropogenic factors influence the distribution and ecological differentiation of *P. brasiliensis* phylogenetic lineages (S1, PS2, and PS3) and *P. lutzii* across South America, the following workflow was implemented:

1) **Occurrence compilation:** Genetically verified occurrence records of *P. brasiliensis* lineages and *P. lutzii* were compiled from published epidemiological and ecological studies. All coordinates were standardized to World Geodetic System 1984 (WGS84), checked for spatial consistency, and matched with environmental layers. Because occurrence data for the *P. brasiliensis* PS4 lineage were sparse, this lineage was not included in the analyses.
2) **Environmental data extraction:** Elevation, degree of urbanisation (Global Human Settlement Layer – Settlement Mode; GHS-SMOD), rainfall erosivity, Köppen–Geiger climate classes, and 19 WorldClim bioclimatic variables were extracted at each occurrence point to generate a harmonized environmental dataset.
3) **Environmental characterization:** Köppen climate classes and Food and Agriculture Organization (FAO) soil orders were assigned to each record to describe the climatic and edaphic context of species and lineage occurrences.
4) **Ordination analysis:** Principal Coordinates Analysis (PCoA) was applied to numeric environmental variables to visualize ecological dissimilarities among *P. brasiliensis* lineages and *P. lutzii*.
5) **Environmental driver analysis**: To identify the environmental gradients underlying the ordination structure, univariate linear regression models were fitted between each environmental variable and the first two PCoA axes.
6) **Cluster identification:** K-means clustering was applied to the environmental dataset to identify environmentally coherent clusters among *P. brasiliensis* occurrences. The optimal number of clusters was determined using the Silhouette Score.
7) **Multivariate cluster comparison:** Convex hulls were constructed in PCoA space for each species and lineage group, and centroid-based hierarchical cluster analysis (HCA) was applied to assess higher-level environmental relationships among lineages.
8) **Decision-tree interpretation:** Decision tree models were fitted to the k-means clusters to identify threshold-based environmental rules underlying cluster separation.
9) **Ensemble machine-learning analyses:** Four ensemble classifiers including Random Forest (RF), Gradient Boosting Machine (GBM), eXtreme Gradient Boosting (XGB), Extra Trees (ET), were used to assess variable importance across pairwise species and lineage comparisons, with performance evaluated via cross-validation and bootstrap resampling.
10) **Climatic suitability modelling:** Climatic suitability was modelled at the species level for *P. brasiliensis* S1 lineage using 19 bioclimatic variables and an ensemble of Random Forest (RF), Gradient Boosting Machine (GBM), eXtreme Gradient Boosting (XGB), and Extra Trees (ET) models. Ensemble predictions were averaged across models to produce a final climatic suitability map with model performance and spatial autocorrelation assessed using the Boyce index and Moran's I statistics.
11) **Ecological differentiation assessment**: The environmental niches and tolerances of *P. brasiliensis* lineages and the xenarthran host *D. novemcinctus* were quantitatively characterized across South America using multivariate analyses and pairwise niche-overlap metrics. By integrating climatic, edaphic, and host-association data, Schoener's D and Hellinger distances were calculated to quantify ecological overlap and dissimilarity among fungal phylogenetic lineages and between lineages and the studied xenarthran host.
12) **SDG-oriented interpretation**: Ecological, climatic, and host-association patterns identified for *Paracoccidioides* species and phylogenetic lineages were interpreted within the framework of the United Nations Sustainable Development Goals (SDGs), with emphasis on goals related to public health, environmental change, biodiversity conservation, and socioeconomic vulnerability.

2.2. Data sources

2.2.1. Occurrence data of Paracoccidioides species and phylogenetic lineages

Georeferenced occurrence records of *P. brasiliensis* (lineages S1, PS2, and PS3) and *P. lutzii* were compiled from published literature. Only occurrences confirmed through genetic or species-level identification were included. Records were extracted from the following key sources: (Hrycyk et al., 2018; Fabris et al., 2024; Bosco and Bagagli, 2018; De Macedo et al., 2019; Mattos et al., 2021; Theodoro et al., 2012). In the case of *P. brasiliensis*, the analyses focused on the major phylogenetic lineages S1, PS2, and PS3 to evaluate lineage-level differentiation. All presence points were standardized to WGS84 and verified for geospatial consistency before environmental extraction.

2.2.2. Xenarthran occurrence data

Occurrence records for the xenarthran species *D. novemcinctus* were obtained from the Global Biodiversity Information Facility (GBIF; URL: www.gbif.org). To ensure compatibility with the environmental dataset and to allow direct comparison with *Paracoccidioides* occurrences, only host records falling within the same Global Human Settlement Layer (GHSL) grid cells that contained *Paracoccidioides* observations were retained for analysis.

The *D. novemcinctus* dataset was downloaded the same day (creation date: 10:52:08), containing 22,734 records from 487 published datasets (compressed size: 1.1 MB): GBIF DOI: https://doi.org/10.15468/dl.avdu86; activation may require several hours). After download, both datasets were cleaned to remove records lacking valid geographic coordinates, duplicates, or obvious spatial errors. Only those GBIF occurrence points intersecting GHSL grid cells that contained *Paracoccidioides* records were used in subsequent environmental and clustering analyses to ensure spatial comparability between pathogen and xenarthran occurrence datasets.

2.2.3. Environmental data

Environmental variables representing terrain, climate, human influence, and rainfall erosivity were compiled as follows:

- **Elevation**: Elevation values were obtained from the ETOPO Global Relief Model (MacFerrin et al., 2024).
- **Rainfall erosivity**: Global rainfall erosivity (R-factor) data were sourced from the European Soil Data Centre (ESDAC), based on the work of Panagos et al. (2022).
- **Climatic data**: Long-term (reference period: 1970–2000) climatic conditions at occurrence sites were derived from the WorldClim 2.1 database (Cerasoli et al., 2022). In addition to standard climate and land-cover variables, aridity was quantified using the Köppen Aridity Index (KAI; Quan et al., 2013).
- **Degree of urbanisation**: Urbanisation categories were derived from the Global Human Settlement Layer – Settlement Model (GHS-SMOD) dataset (Pesaresi et al., 2024) provided by the Joint Research Centre of the European Commission.
- **Climate zones**: Köppen–Geiger climate classifications at 1-km resolution were taken from Beck et al. (2018), derived from WorldClim v2.1 (Fick and Hijmans, 2017).
- **Bioclimatic variables**: Nineteen bioclimatic variables (bio1–bio19), representing long-term averages for 1970–2000, were extracted from WorldClim v2.1 (Fick and Hijmans, 2017). These variables capture temperature and precipitation means, seasonality, and extremes.

All raster datasets were processed in Quantum Geographic Information System (QGIS) 3.34.1 in combination with GRASS GIS 8.4.0 software, where point values for each environmental layer were sampled at species occurrence coordinates.

Variables used in the analyses are given in **Supplementary Table S1**.

2.3. Ordination of environmental space (PCoA)

Environmental similarity among occurrence sites was assessed using Principal Coordinates Analysis (PCoA). A dissimilarity matrix was computed from standardized environmental variables (Gower, 2014) using Euclidean distance, after scaling variables to zero mean and unit variance to ensure comparability across different measurement units. PCoA was then applied to the distance matrix to reduce the multivariate environmental space into a lower-dimensional ordination framework, allowing visualization of major ecological gradients.

2.4. Cluster delineation: K-means and convex-hull–based HCA

To identify ecologically coherent groups independent of taxonomy, k-means clustering was applied to the environmental dataset. The optimal number of clusters (k) was determined with the Silhouette Score, maximizing both internal compactness and separation (Saputra et al., 2020). Spatial relationships among clusters were examined with hierarchical clustering analysis (HCA) performed on the centroids of PCoA-derived convex hulls, providing an additional multivariate structure summary.

2.5. Environmental driver analysis (ordination interpretation)

To interpret the ecological gradients underlying the ordination structure, univariate linear regression models were fitted between each environmental variable and the first two principal coordinates (PCoA1 and PCoA2). For each variable, the coefficient of determination (R^2) and associated F-statistics were computed to quantify the strength of association between environmental predictors and ordination axes. The combined explanatory power (total R^2) was used to rank variables according to their overall contribution to environmental differentiation.

2.6. Decision-tree classification of environmental clusters

The environmental clusters identified through k-means were interpreted and classified via supervised decision-tree analysis. The decision trees identified threshold-based environmental rules distinguishing cluster membership. Internal nodes represented decision splits, and leaves corresponded to final cluster assignments.

2.7. Ensemble machine-learning models for feature importance

To assess environmental determinants distinguishing species and phylogenetic lineages, four ensemble learning models were applied:

- Random Forest (RF)
- Gradient Boosting Machine (GBM)
- eXtreme Gradient Boosting (XGBoost)
- Extra Trees (ET)

Pairwise classification was performed for: (a) *P. brasiliensis* vs. *P. lutzii*, (b) *P. lutzii* vs. *P. brasiliensis* PS2, (c) *P. brasiliensis* S1 vs. PS2+PS3, (d) *P. brasiliensis* S1 vs. PS2.

Models were trained with an 80/20 train–test split and evaluated using accuracy, F1-score, RMSE, and R^2, combined with 5-fold cross-validation and 100 bootstrap resampling iterations. Each model produced feature-importance scores, which were compared to identify consistent predictors across analyses.

2.8. Climatic suitability modelling

Climatic suitability was modelled for the *P. brasiliensis* S1 lineage using an ensemble machine-learning framework combining multiple tree-based algorithms and presence–background data. Nineteen bioclimatic variables (bio1–bio19) representing long-term temperature and precipitation patterns were used as predictors. Values were extracted from raster layers at species occurrence locations and randomly generated background points using bilinear interpolation.

Occurrence records of the *P. brasiliensis* S1 lineage were obtained from a curated database of genetically confirmed sites and cleaned to remove incomplete or inconsistent entries. Background points were generated randomly within a predefined South American extent (longitude: −90° to −30°, latitude: −35° to 15°) with a minimum distance of 300 km from any known occurrence to reduce spatial bias and autocorrelation, retaining 5,000 points. Presence points were assigned a value of 1, background points 0, and records with missing data were excluded. The dataset was split into training (80%) and testing (20%) subsets using a fixed random seed.

Suitability was estimated using an ensemble of Random Forest,

Gradient Boosting Machine, eXtreme Gradient Boosting, and Extra Trees classifiers. Each model produced probabilistic predictions, and final ensemble scores were calculated as the mean probability across all four models.

Mathematically, for any location x:

$$Ensemble\ suitability(x) = \frac{P_{RF}(x) + P_{GBM}(x) + P_{XGB}(x) + P_{ET}(x)}{4}$$

Model performance was assessed using the Boyce index, a presence-only metric suitable for ecological niche and suitability modelling. The index was calculated by comparing ensemble-predicted suitability at presence locations against background points using quantile-based binning, with the final value given by the Spearman rank correlation between suitability class midpoints and observed-to-expected presence ratios (Zhang and Wang, 2023; Liu et al., 2025).

Spatial predictions were generated across the entire raster domain using a block-wise approach to optimise memory. Bioclimatic layers were processed in fixed-size blocks, and ensemble suitability values for each cell were calculated as the mean of probabilistic predictions across models. The final climatic suitability map was exported as a single-band floating-point Geographic Tagged Image File Format (GeoTIFF) with Lempel–Ziv–Welch (LZW) compression.

Model robustness was further evaluated by Moran's I to quantify spatial autocorrelation in occurrence data across neighbour distances from 10 km to 1000 km, ensuring that predicted suitability reflected genuine environmental associations rather than clustered sampling artifacts.

All analyses were performed in Python 3.10 using standard geospatial and machine-learning libraries, with reproducibility ensured by a fixed random seed. Raster outputs were visualised in QGIS 3.34.11 and GRASS GIS 8.4.0.

2.9. Quantitative niche overlap among Paracoccidioides lineages and xenarthran hosts

Environmental niche overlap among *Paracoccidioides* lineages and xenarthran hosts was quantified using Schoener's D and Hellinger distance across a comprehensive set of environmental variables, including Köppen climate class, 19 Bioclimatic variable dataset (BIOCLIM) variables, elevation, global erosion, ruggedness, soil type, Köppen Aridity Index (KAI), and degree of urbanisation (GHSL).

For each pair of lineages or host species, occurrence points were used to extract environmental values from the full predictor set. Schoener's D measures niche overlap from 0 (no overlap) to 1 (complete overlap), while Hellinger distance quantifies dissimilarity from 0 (identical niches) to 1 (maximally distinct). Continuous variables were normalized, and categorical variables (e.g., climate and soil type) were one-hot encoded to allow inclusion in the multivariate analysis. Pairwise metrics were computed for all lineage–lineage, lineage–host, and host–host comparisons, producing a quantitative matrix of niche similarity and dissimilarity.

2.10. SDG-oriented interpretation framework

To place the ecological and epidemiological findings into a broader sustainability context, results were interpreted according to the United Nations Sustainable Development Goals (SDGs). Associations between *Paracoccidioides* species or phylogenetic lineages and specific SDGs were evaluated qualitatively based on observed relationships among climatic suitability, environmental specialisation, land-use patterns, host associations, and potential disease-risk implications.

The interpretation particularly considered SDG 3 (Good Health and Well-being), SDG 6 (Clean Water and Sanitation), SDG 8 (Decent Work and Economic Growth), SDG 10 (Reduced Inequalities), SDG 13 (Climate Action), and SDG 15 (Life on Land). SDG assignments were based on the ecological characteristics of each lineage, including climatic sensitivity, habitat specialisation, geographic restriction, environmental overlap with xenarthran hosts, and potential vulnerability to environmental and anthropogenic change.

This framework was intended as an integrative ecological interpretation rather than a quantitative SDG assessment, allowing the ecological differentiation and climatic vulnerability of *Paracoccidioides* lineages to be linked to broader environmental-health and sustainability challenges across South America.

3. Results

3.1. Environmental characterization

In the elevation map (Fig. 1**a**), most sampling locations are concentrated within the Brazilian states of Mato Grosso, Goiás, Minas Gerais, São Paulo, and Mato Grosso do Sul, regions dominated by mid-elevation uplands. Overall, occurrences cluster around ~560 m above sea level (mean: 558 ± 218 m; range: 39–853 m), suggesting that this altitudinal zone provides suitable environmental conditions for all four examined taxa.

Against this general elevational background, lineage-specific patterns emerge. *Paracoccidioides brasiliensis* S1 shows the broadest elevational distribution. Although most S1 occurrences are associated with mid-elevation terrains, several records originate from the Amazon Basin at lower elevations, resulting in a wide overall range (mean: ~540 m; SD: 212 m; range: 39–853 m). This pattern indicates that the S1 subtype occupies both lowland tropical environments and higher-elevation habitats.

In contrast, *P. brasiliensis* PS2 and PS3, as well as *P. lutzii*, exhibit more restricted elevational ranges and are predominantly associated with mid-elevation regions. PS2 occurrences are centered around ~670 m (mean: 667 ± 214 m; range: 113–919 m), whereas PS3 and *P. lutzii* are typically found near ~540 m above sea level (mean: 540 ± 209 m; range: 168–772 m).

The Köppen–Geiger climate categories show variation among taxa. *Paracoccidioides brasiliensis* (S1, PS2, and PS3) occurs primarily in tropical monsoon (Köppen Am), tropical savannah (Köppen Aw), and hot summer humid subtropical temperate (Köppen Cfa) climate zones, with fewer records in tropical rainforest (Köppen Af) and warm summer humid subtropical temperate (Köppen Cfb) zones (Fig. 1**b**; Fig. 2**a1**). Among lineages, S1 is the only group recorded in Af and Am climates. PS2, PS3, and *P. lutzii* are mainly associated with Aw and Cfa climates, while only PS2 is also observed in Cfb zones (Fig. 2**b1**). Overall, tropical savannah (Aw) represents the dominant climatic environment supporting all examined lineages, whereas Cfa also appears suitable for the more widespread lineages S1 and PS2. However, these patterns should be interpreted with caution, as the observed distributions reflect available literature and database records and may be influenced by spatial and reporting bias associated with underdiagnosis and uneven sampling effort, rather than the full ecological range of each lineage.

When occurrences were analysed in relation to FAO soil types, *P. brasiliensis* was predominantly associated with oxisols (udox and ustox) and ultisols (udult) (Fig. 2**a2**). In addition, it occurred sporadically across a wider range of less frequent soil categories, including alfisols (udalf and ustalf), aridisols (calcid), entisols (psamment and orthent), and inceptisols (aquept and udept), as well as perox oxisols and ustult ultisols, indicating a relatively broad but uneven edaphic distribution.

Paracoccidioides lutzii showed partial overlap with *P. brasiliensis*, particularly in oxisol-dominated regions (udox and ustox). However, it exhibited its strongest association with ustult ultisols (Fig. 2b2). In contrast, *P. brasiliensis* was only rarely recorded in ustult ultisols environments. Notably, *P. lutzii* was not observed across the other soil categories recorded for *P. brasiliensis* suggesting a more restricted edaphic distribution. Given the soil-dwelling nature of these fungi, these patterns are consistent with a potential influence of soil properties on their

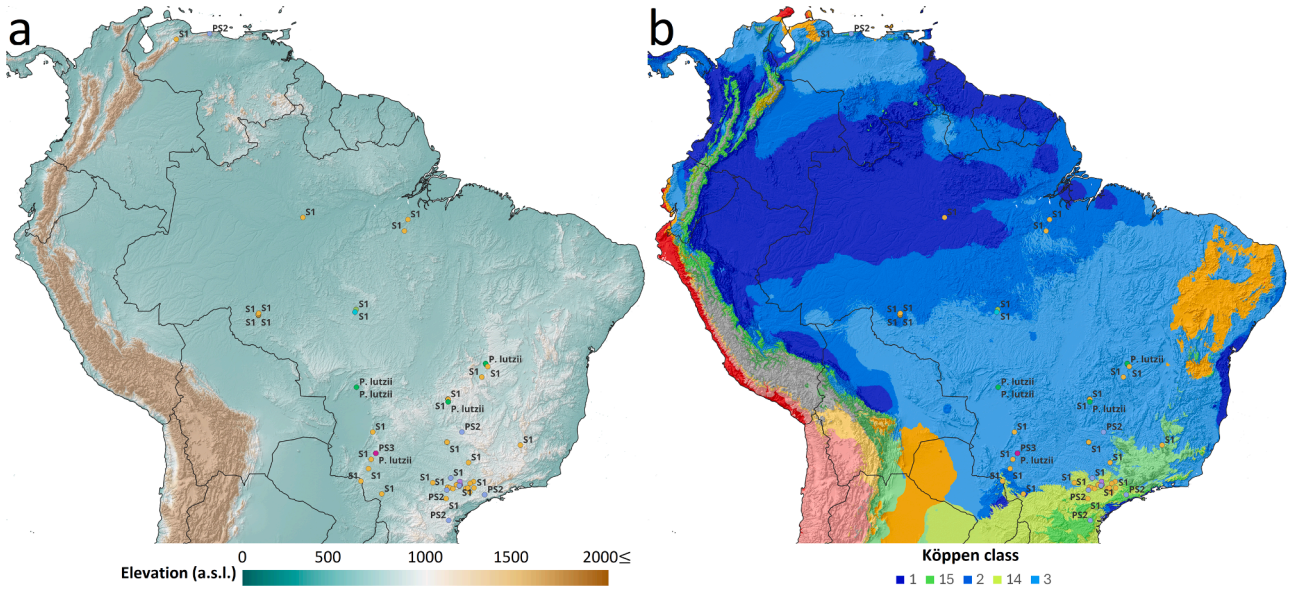

Fig. 1. The occurrences of different lineages of *Paracoccidioides braziliensis* and *Paracoccidioides lutzii* projected onto the elevation (a) and Köppen (b) maps of northern and central South America. Within *P. brasiliensis*, S1 (sensu stricto), PS2, PS3, and PS4 represent phylogenetic lineages of the species. Köppen climate classes are provided in **Supplementary Table S1**. Occurrence records are derived from literature and databases and may be affected by spatial and reporting bias due to underdiagnosis, thus representing known records rather than the full geographic distribution.

ecological distribution, although they should be interpreted considering sampling limitations and uneven occurrence reporting.

3.2. PCoA and HCA results

The PCoA ordination based on standardized environmental variables (Fig. 3a) did not reveal discrete clustering among the analysed lineages. Instead, the ordination space exhibited a largely continuous structure with substantial overlap among *P. lutzii*, PS2, PS3, and S1, indicating a gradient-like organization of environmental conditions rather than clearly separated ecological clusters.

The first two PCoA axes captured the dominant environmental gradients, primarily driven by temperature-related variables (bio11, bio6, bio1) along PC1, and precipitation- and aridity-related variables (bio12, bio13–bio18, KAI) along PC2. However, visual inspection of the ordination confirmed that group separation is partial and continuous rather than discrete.

Quantitatively, centroid-based analyses supported this pattern. Although between-group distances exceeded within-group dispersion (global separation index = 1.87), overall clustering quality was low, as indicated by a negative silhouette score (−0.39). This further confirms that the observed structure does not correspond to well-defined, compact clusters in multivariate space.

Within-group dispersion showed marked heterogeneity, with PS3 exhibiting near-zero internal variation (0.00), likely reflecting limited sample size or extreme environmental constraint, while S1 showed the highest dispersion (3.44), consistent with a broad ecological tolerance. *P. lutzii* and PS2 exhibited intermediate variability (2.24 and 2.67, respectively).

Centroid distances revealed partial ecological differentiation without clear partitioning. The greatest separation was observed between *P. lutzii* and PS2 (6.98), whereas the smallest distance occurred between PS3 and S1 (1.82). Importantly, *P. lutzii* occupied an intermediate position in ordination space, showing moderate distances to both PS3 (3.23) and S1 (2.89), reinforcing the absence of sharp ecological boundaries.

In summary, the PCoA results indicate a continuum of environmental niches rather than discrete clustering, with PS2 partially differentiated from the remaining lineages, while PS3 and S1 occupy overlapping regions of environmental space. *Paracoccidioides lutzii* appears as an intermediate lineage within this gradient structure.

The centroid-based hierarchical clustering (Fig. 3b) partially reflects the ordination structure, grouping S1 and PS3 as the closest pair, while *P. lutzii* occupies an intermediate position between this cluster and PS2. PS2 remains the most environmentally distinct lineage, although this separation is not absolute but gradient-based rather than cluster-based.

3.3. Environmental drivers of ordination structure

The heatmap analysis of environmental drivers (Fig. 4) revealed a strong and heterogeneous influence of climatic variables on the ordination structure. Rather than a single dominant predictor, multiple environmental factors jointly contributed to variation along the first two PCoA axes, supporting a multidimensional environmental gradient.

The strongest associations were observed for temperature-related variables. In particular, bio11 (mean temperature of coldest quarter; $R^2 = 0.98$) and bio6 (minimum temperature of coldest month; $R^2 = 0.95$) exhibited the highest explanatory power for PC1, followed closely by bio1 (annual mean temperature; $R^2 = 0.95$). These results indicate that low-temperature constraints represent a primary axis of environmental differentiation.

In contrast, PC2 was primarily driven by precipitation and water-balance variables. The variables bio12 (annual precipitation; $R^2 = 0.90$) and bioclimatic moisture-related variables (bio13–bio18) showed moderate to strong contributions, indicating that hydrological gradients structure the second major axis of environmental variation.

Additionally, the aridity index (KAI; $R^2 = 0.90$) emerged as a key secondary driver, reinforcing the importance of water limitation in shaping environmental space. Variables such as elevation, ruggedness, and some derived bioclimatic indices showed weaker but still measurable effects, suggesting a secondary role in fine-scale environmental structuring.

K-means clustering applied to the environmental dataset, visualised in PCA space (Fig. 5), identified seven environmentally coherent clusters within *P. brasiliensis* occurrences (**Supplementary Figure S1**). The clustering structure showed moderate but consistent separation (silhouette score = 0.289), supported by additional validation metrics including a Calinski–Harabasz index of 20.85 and a Davies–Bouldin

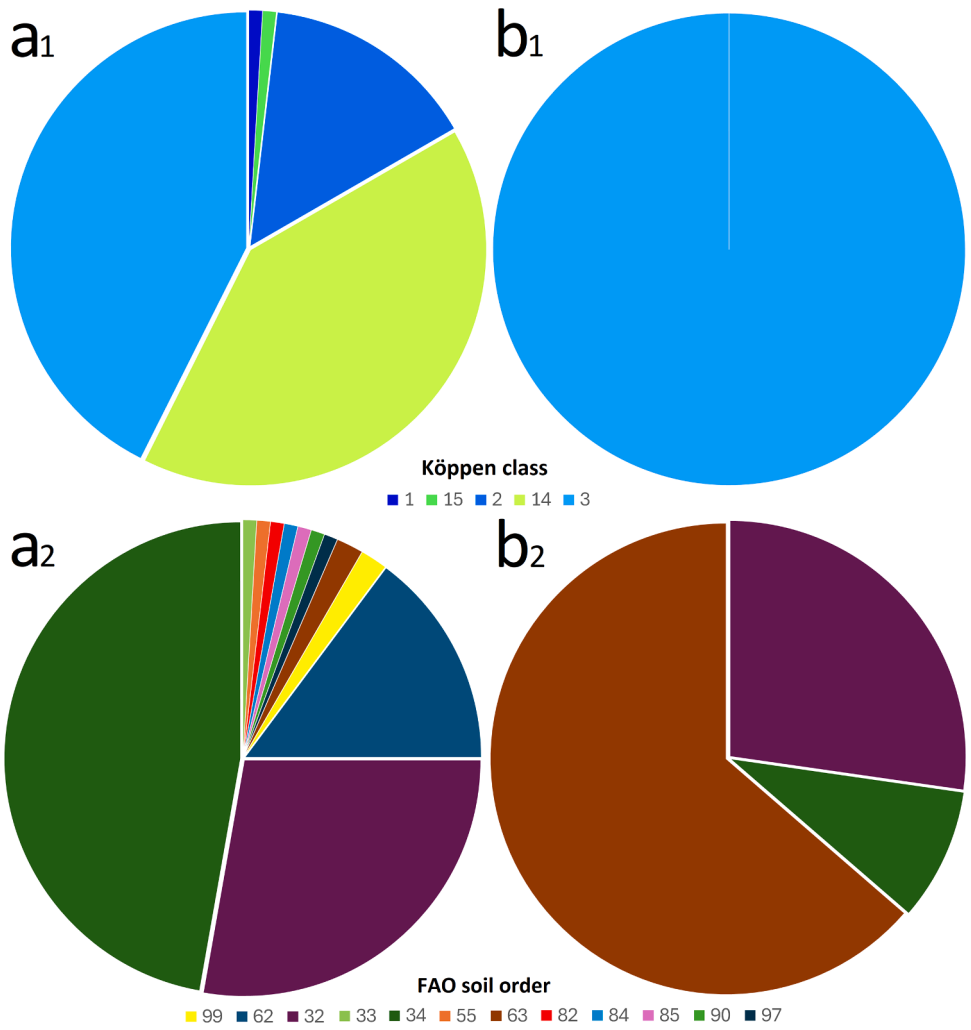

Fig. 2. Proportional representation of Köppen climatic classes (1) and FAO soil orders (2) among occurrences of *Paracoccidioides braziliensis* (a) and *Paracoccidioides lutzii* (b). Köppen climate classes and FAO soil categories are provided in **Supplementary Table S1**.

index of 1.48, indicating distinguishable yet partially overlapping ecological niches.

The PCA biplot revealed that the clustering structure is primarily structured by temperature and energy-related variables together with topographic and edaphic factors. Mean annual temperature (bio1), mean temperature of coldest quarter (bio11), elevation, and udult ultisols (soil grid code 62) emerged as the dominant environmental gradients shaping cluster separation, with additional contributions from seasonal temperature variability (bio15) and precipitation seasonality (bio8, bio19) playing secondary roles.

Clusters located in the upper portion of the ordination space are strongly associated with lower temperatures and higher elevations, indicating cooler ecological conditions, whereas clusters positioned in the opposite direction reflect warmer lowland environments. Among these, clusters characterised by elevated values of bio11, and elevation form the most distinct ecological groupings, suggesting a strong altitudinal control on niche differentiation. Soil type (particularly udult ultisols) further refines this structure, indicating that edaphic constraints contribute substantially to within-temperature niche partitioning.

3.4. Decision tree analyses results

Decision tree analyses were subsequently applied to interpret the environmental rules underlying the k-means-derived clusters (**Supplementary Fig. S2**). The decision tree analysis confirmed that cluster separation is predominantly controlled by a small subset of environmental predictors. Mean annual temperature (bio11) emerged as the most important variable (33.0%), followed by udult ultisols (soil grid code 62; 23.7%) and elevation (20.9%). These three variables jointly explain the majority of classification structure, highlighting a strong temperature–topography–soil axis underlying ecological differentiation.

Seasonal precipitation and temperature variability (bio15, bio8, bio19) contributed only marginally to cluster discrimination, while several variables—including bio1 and multiple soil categories—showed no measurable contribution in the tree-based partitioning. The decision tree further revealed that clusters form hierarchically structured decision boundaries primarily driven by temperature thresholds and elevation cut-offs, with soil type acting as a secondary but still important splitting criterion. This indicates that while k-means captures multivariate ecological structure, the underlying separations are largely governed by a small number of dominant environmental constraints rather than broad climatic gradients. The complete decision tree structure can be found in **Supplementary Text S1**.

3.5. Machine learning ensemble analyses

Machine learning ensemble analyses were applied to quantify the extent to which environmental variables discriminate among *Paracoccidioides* lineages and to identify the predictors contributing most

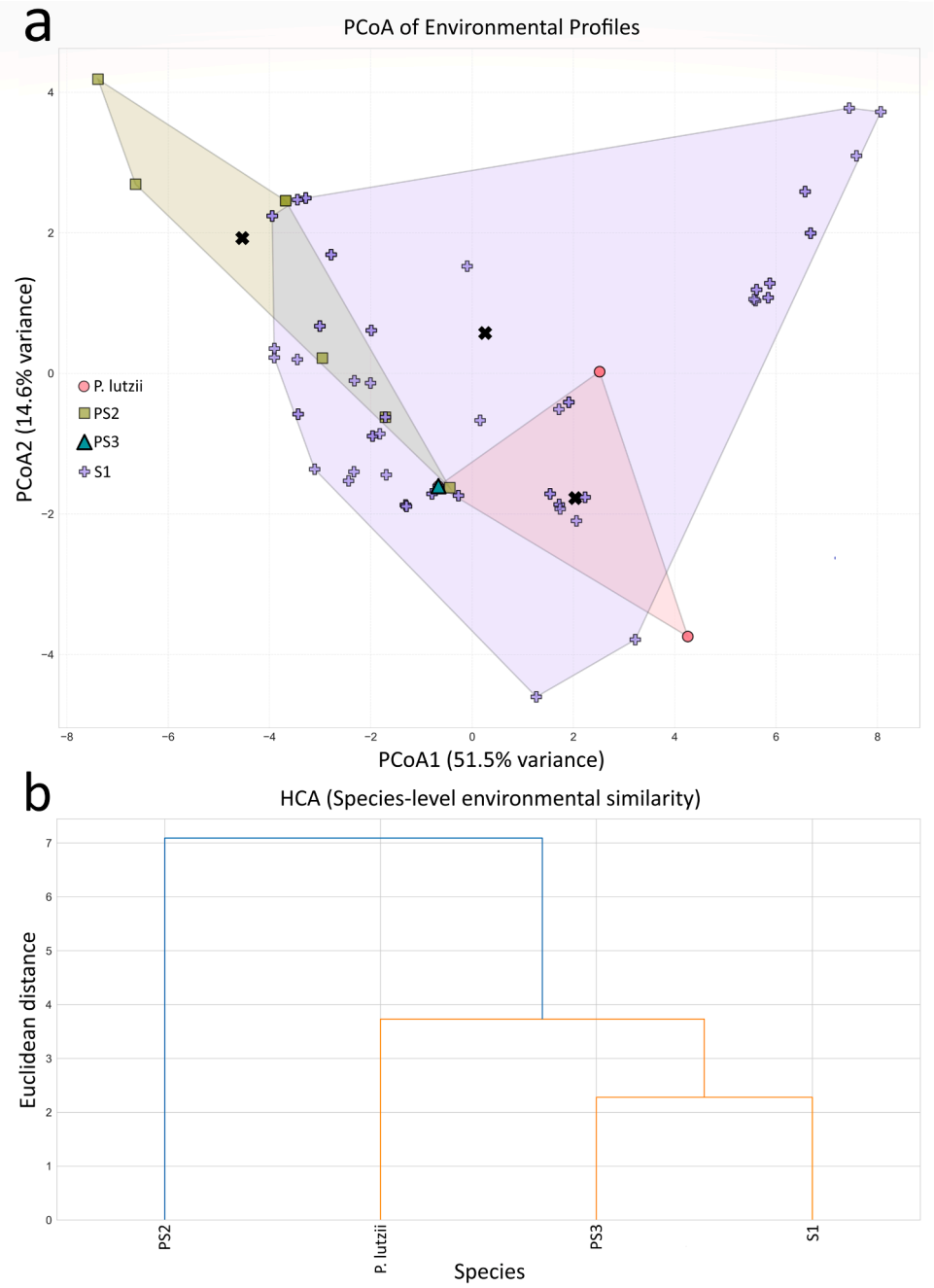

Fig. 3. PCoA analysis of environmental variables associated with *Paracoccidioides* lineages. **a**: Principal Coordinate Analysis (PCoA) based on Euclidean distances of standardized environmental variables, showing convex hulls for each lineage. **b**: Centroid-based hierarchical clustering derived from PCoA coordinates, illustrating relative environmental similarity among lineages.

strongly to classification performance (Fig. 6a–b). Rather than representing a direct measure of ecological distance, differences in model performance reflect the degree of separability of the groups within multivariate environmental space.

The comparison between *P. brasiliensis* and *P. lutzii* (Fig. 6a) yielded consistently very high predictive performance across all ensemble approaches, with classification accuracies ranging from 0.966 (XGBoost) to 0.992 (Extra Trees). The uniformly high F1-scores, low prediction errors, and moderate-to-high agreement among models indicate a highly robust separation between the two taxa, suggesting that they occupy markedly distinct environmental niches. Feature importance analyses consistently identified temperature- and precipitation-related variables as major discriminators, particularly maximum temperature of the warmest month (bio5), mean temperature of the warmest quarter (bio10), precipitation of the driest quarter (bio17), and precipitation of the coldest quarter (bio19). Additional contributions from precipitation of the driest month (bio14), precipitation seasonality (bio15), annual mean temperature (bio1), and landscape-related predictors such as ruggedness and GHSL-derived variables further suggest that both climatic and environmental heterogeneity contribute to species-level differentiation.

In contrast, the comparison between *P. lutzii* and *P. brasiliensis* lineage PS2 (Fig. 6**b**) resulted in more variable model performance, with accuracies ranging from 0.864 (Random Forest) to 0.955 (XGBoost and Voting Ensemble). Although the groups remain distinguishable based on environmental predictors, the broader variation among classifiers and

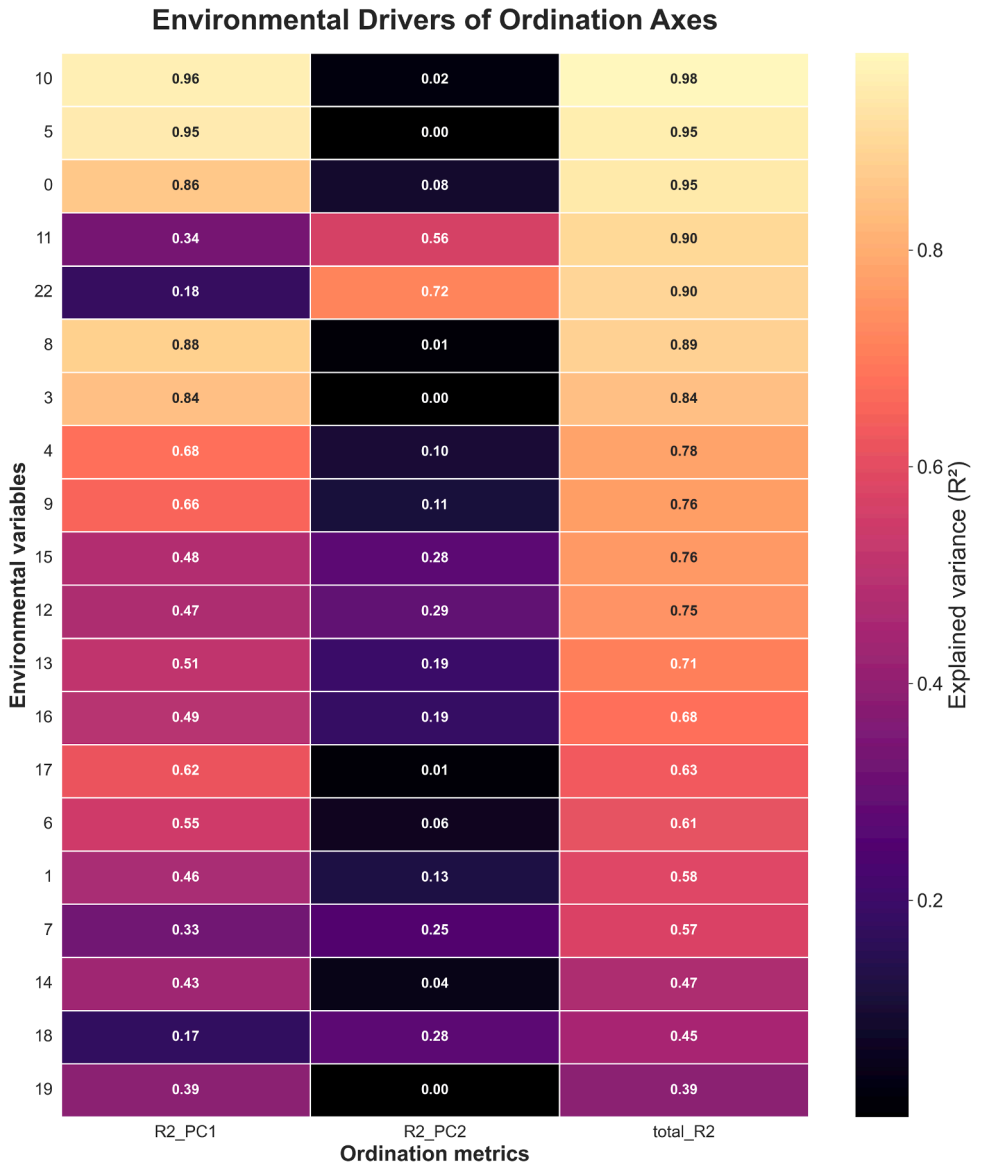

Fig. 4. Environmental drivers of ordination structure. Heatmap showing the explanatory power (R^2 values) of environmental variables in relation to the first two PCoA axes and their combined effect (total R^2). Variables are ranked according to overall explanatory strength. The k-means clustering results.

lower agreement between models suggest a less sharply defined and more model-dependent ecological separation relative to the species-level comparison.

In this case, classification was primarily driven by annual mean temperature (bio1), maximum temperature of the warmest month (bio5), mean temperature of the warmest quarter (bio10), and several precipitation-related variables, including bio14, bio15, bio17, and bio19.

In this case, classification was primarily driven by annual mean temperature (bio1), maximum temperature of the warmest month (bio5), mean temperature of the warmest quarter (bio10), and several precipitation-related variables, including precipitation of driest month (bio14), precipitation seasonality (bio15), precipitation of driest quarter (bio17), and precipitation of coldest quarter (bio19). Together, these results indicate that thermal and moisture regimes remain central drivers of ecological differentiation within the *P. brasiliensis* complex, although the environmental boundaries between *P. lutzii* and PS2 appear comparatively less distinct.

Table 1 shows the model performance metrics of the applied ensemble models, and the complete correlation matrices among all environmental predictors are provided in **Supplementary Figure S3**, illustrating the interdependencies and potential redundancies within the full predictor set.

3.6. *Environmental partitioning among Paracoccidioides lineages and xenarthran hosts quantified by Schoener's D and Hellinger distance*

Pairwise environmental niche overlap among *Paracoccidioides* lineages and their xenarthran hosts was quantified using Schoener's D (Fig. 7a) and Hellinger distance (Fig. 7b) across 26 environmental predictors, including climatic, edaphic, and anthropogenic variables. Values of Schoener's D range from 0 (no overlap) to 1 (complete overlap), while Hellinger distance measures dissimilarity from 0 (identical) to 1 (maximally different). The analysis revealed that the xenarthran host *D. novemcinctus* occupies a broad and heterogeneous environmental space, exhibiting high overlap with all fungal lineages (D = 0.91–0.97; Hellinger = 0.036–0.091). Similarly, *P. brasiliensis* S1 shows substantial overlap with other fungal lineages, particularly PS3 (D = 0.95; H = 0.047), reflecting its broad ecological tolerance and flexible habitat use. In contrast, *P. lutzii* exhibits lower overlap with other lineages (D =

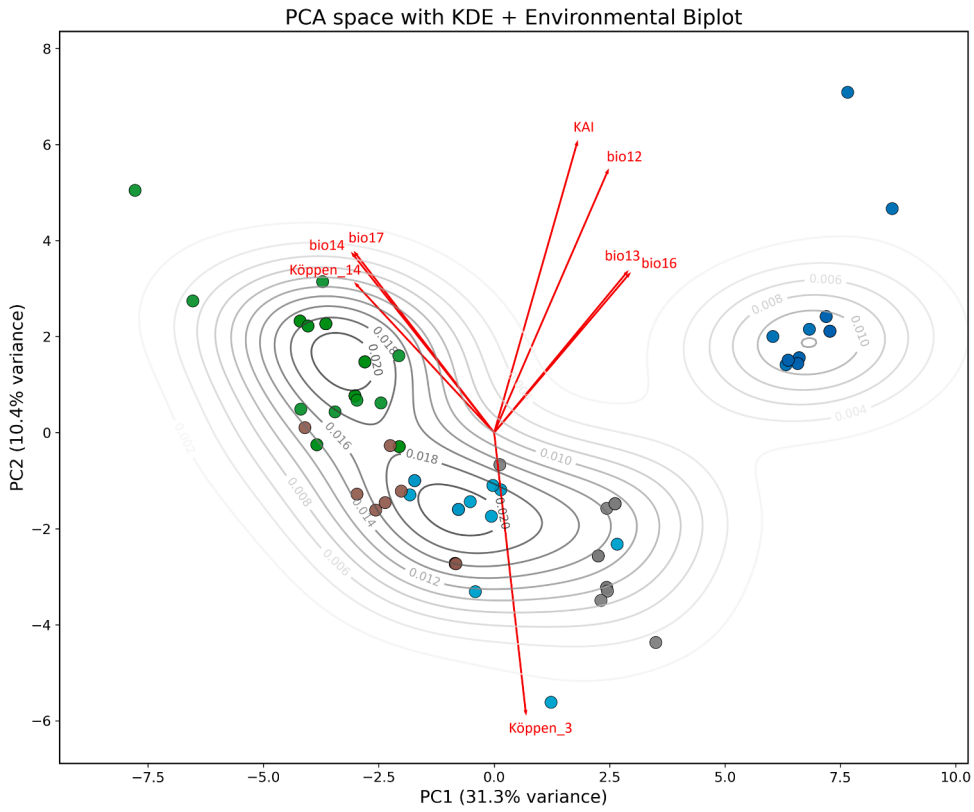

Fig. 5. K-means clustering results of the environmental characteristics of all the studied genetic lineages of *Paracoccidioides braziliensis* in PCA space. Colored points represent the seven environmental clusters identified by k-means analysis, contour lines indicate kernel density estimation (KDE) of occurrence density, and red vectors show the environmental variables with the strongest contributions to the first two principal components. Bioclimatic classes are given in **Supplementary Table S1**.

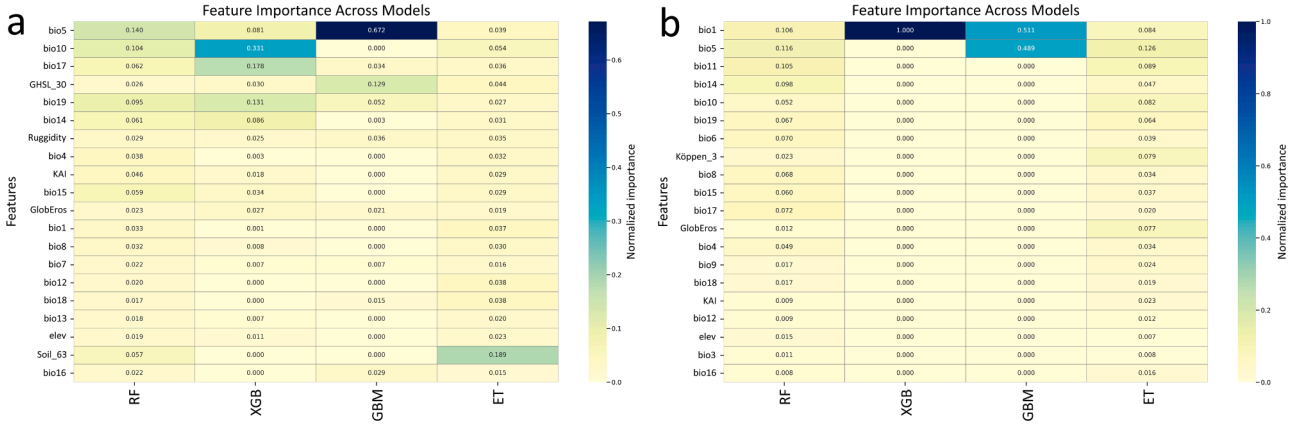

Fig. 6. Feature importance comparisons of the occurrence-determining environmental conditions of different *Paracoccidioides* lineages using Gradient Boosting Machine (GBM), Random Forest (RF), eXtreme Gradient Boosting (XGB), and Extra Trees (ET) machine learning methods. **a**: *Paracoccidioides braziliensis* versus *Paracoccidioides lutzii*, **b**: *Paracoccidioides lutzii* versus *Paracoccidioides braziliensis* lineage PS2. The explanation of the environmental factors can be found in **Supplementary Table S1**.

0.87–0.94; H = 0.069–0.122), consistent with its narrower geographic distribution and specialized environmental requirements.

3.7. Modelled climatic suitability patterns

The Random Forest–based climatic suitability model for *P. brasiliensis* S1 main lineage reveals a well-defined spatial structure, with high suitability concentrated primarily in southeastern Brazil (Fig. 8). The core area of suitability is spatially coherent but comparatively compact, encompassing parts of Minas Gerais, São Paulo, Goiás, and Mato Grosso do Sul. Climatic suitability declines more rapidly toward the margins of this core, resulting in relatively narrow transitional zones and indicating stronger environmental constraints at the edges of the distribution. This pattern suggests a moderately broad but environmentally delimited climatic niche, with persistence favoured under a specific combination of climatic conditions rather than across uniformly permissive environments. Model robustness is supported by a high Boyce index (0.866; ensemble, M-area), while near-zero Moran's I value across relevant neighbour distances indicate negligible spatial autocorrelation among occurrence records, reinforcing the reliability of the

Table 1
Model performance metrics based on 5-fold cross-validation. Accuracy and F1-score reflect classification performance, while RMSE is computed from label-encoded predictions and should be interpreted as a distance-based classification error rather than a regression metric. Differences in performance between comparisons reflect variation in the intrinsic separability of environmental predictors among lineages.

Comparison	Modell	Accuracy	F1-score	RMSE	R^2
P. brasiliensis vs *P. lutzii*	Random Forest	0.983	0.983	0.130	0.800
	XGBoost	0.966	0.966	0.183	0.599
	Gradient Boosting	0.983	0.982	0.130	0.800
	Extra Trees	0.992	0.991	0.092	0.900
P. lutzii vs *P. brasiliensis* PS2	Random Forest	0.864	0.863	0.369	0.455
	XGBoost	0.955	0.954	0.213	0.818
	Gradient Boosting	0.909	0.909	0.302	0.636
	Extra Trees	0.909	0.909	0.302	0.636

inferred spatial patterns (**Supplementary Figure S4**).

3.8. SDG-oriented conclusions by Paracoccidioides species and genetic lineages

Distinct ecological patterns exist among *Paracoccidioides* species and lineages, reflecting differences in environmental breadth, climatic tolerance, and disease-risk potential. The *P. brasiliensis* S1 lineage shows the broadest niche, spanning multiple elevations, five Köppen–Geiger climate zones, and various oxisol and ultisol soils. It overlaps with xenarthran hosts, indicating a generalist strategy. Its widespread soil association links it to SDG 3 (Good Health and Well-being), SDG 8 (Decent Work and Economic Growth), SDG 15 (Life on Land), SDG 13 (Climate Action) and indirectly to SDG 6 (Clean Water and Sanitation). Control efforts require a One Health approach, targeting monitoring, diagnostics, therapy, and prevention.

The *P. brasiliensis* PS2 lineage is highly specialised, restricted to narrow climatic and elevational ranges. Its localised distribution necessitates targeted monitoring and interventions, connecting it to SDGs 3, 8, 10 (Reduced Inequalities), 13, and 15. Local ecosystem changes strongly influence its distribution, making it a valuable indicator for biodiversity and climate impacts. The *P. brasiliensis* PS3 lineage occupies an intermediate niche with partial overlap with S1, showing moderate climatic sensitivity. Control measures relate to SDG 3 and SDG 13, as climate change may expand its distribution.

Paracoccidioides lutzii is highly specialised, confined to a single climate zone, narrow elevation, and udult ultisols, making it sensitive to climatic and land-use changes. It aligns strongly with SDG 3, 13, and 15, serving as a model for studying pathogen responses to environmental change.

Overall, the *Paracoccidioides* complex shows substantial lineage- and species-level ecological divergence, with distinct niche breadth, environmental specialisation, and spatial climatic suitability across South America (Fig. 9).

4. Discussion

This study provides a comprehensive assessment of the environmental, climatic, edaphic, and anthropogenic factors shaping the distribution and ecological differentiation of *P. brasiliensis* genetic lineages (S1, PS2, PS3) and *P. lutzii* across South America. By integrating occurrence records with a broad suite of environmental variables, our analyses indicate that ecological differentiation within the *Paracoccidioides* complex is best interpreted as a predominantly continuous, multidimensional environmental gradient, with lineage-level structuring superimposed on this continuous variation.

The results show that most sampling points are concentrated in mid-elevation uplands (~1,000 m) across Mato Grosso, Goiás, Minas Gerais, São Paulo, and Mato Grosso do Sul, regions long recognised as endemic for *Paracoccidioides* spp. (Arantes et al., 2016). These landscapes provide broadly suitable climatic and ecological conditions for multiple lineages. Among them, S1 exhibits the broadest elevational and geographic distribution, extending into Amazonian lowlands below 500 m, consistent with substantial ecological flexibility and tolerance of both humid lowland environments and seasonally dry uplands (Mattos et al., 2021). In contrast, PS2, PS3, and *P. lutzii* occur predominantly within mid-elevation environments, suggesting comparatively narrower environmental tolerances, although occasional lowland detections indicate

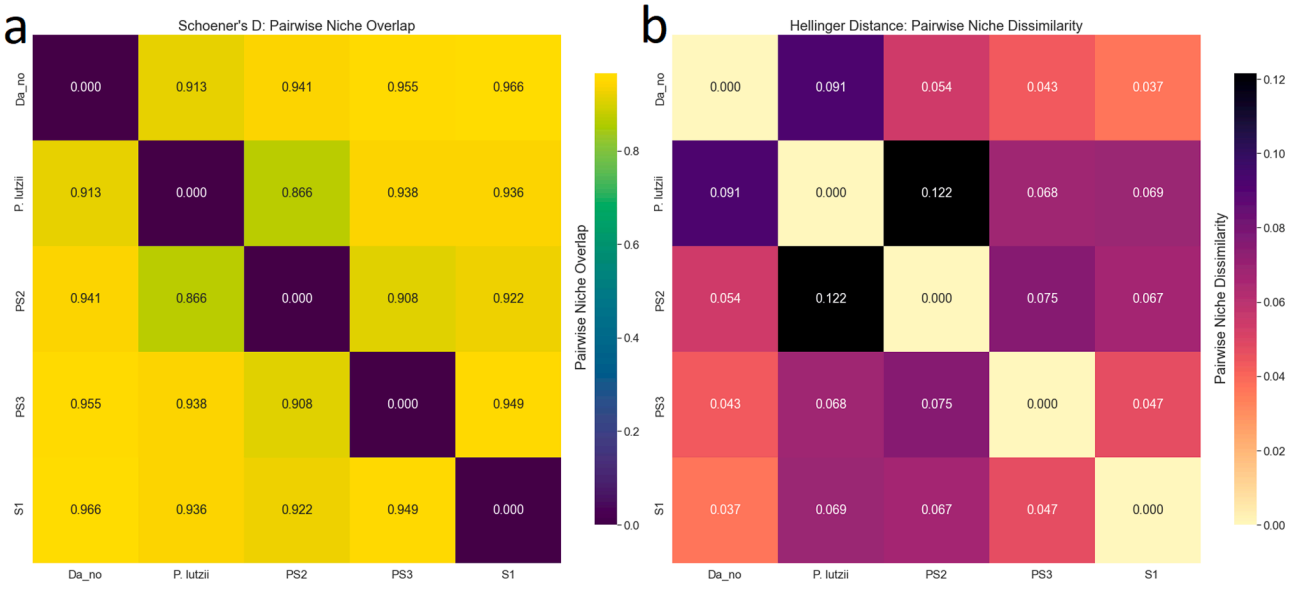

Fig. 7. Pairwise environmental niche comparison among *Paracoccidioides* lineages and xenarthran hosts. Heatmaps show Schoener's D (**a**), representing niche overlap, and Hellinger distance (**b**), representing niche dissimilarity, based on normalized environmental variables (climatic, edaphic, and anthropogenic). Higher Schoener's D values indicate greater similarity, whereas higher Hellinger values indicate stronger ecological divergence. Abbreviations: *P. lutzii: Paracoccidioides lutzii*; Da_no: *Dasypus novemcinctus*; PS2, PS3, and S1: *Paracoccidioides brasiliensis* lineages.

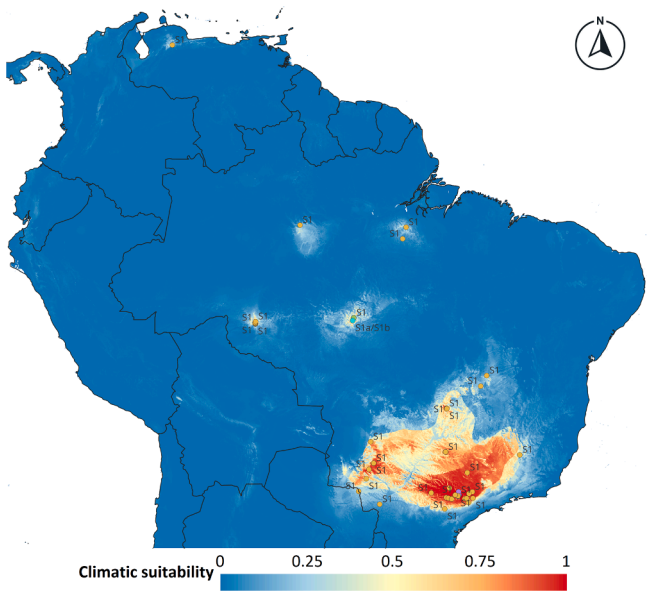

Fig. 8. Modelled climatic suitability for *Paracoccidioides brasiliensis* S1 lineage.

that local microhabitats, soil properties, and reservoir-host ecology may expand realised distributions beyond macroclimatic expectations (Martinez, 2015).

Climatic analyses further support these patterns. Mid-elevation uplands in central and southeastern Brazil provide favourable environmental conditions for *Paracoccidioides* spp., consistent with previous ecological and epidemiological studies (Fabris et al., 2024). S1 spans five Köppen–Geiger climate zones, highlighting broad ecological tolerance, whereas *P. lutzii* is largely confined to tropical savannah climates (Köppen Aw), reflecting more specialized environmental requirements and its predominance in central-western Brazil (Hahn et al., 2019). Tropical savannah (Köppen Aw) and humid subtropical climates (Köppen Cfa) support multiple lineages, emphasising the importance of specific thermal and hydrological regimes in fungal persistence. Nevertheless, occasional detections outside predicted zones suggest that microhabitat heterogeneity, local dispersal processes, host dynamics, and sampling bias also shape realised distributions (Hrycyk et al., 2018).

Soil characteristics emerged as critical determinants of ecological suitability. *Paracoccidioides brasiliensis* occurs across diverse soil classes, especially ustox oxisols, udox oxisols, and udult ultisols, whereas *P. lutzii* shows stronger association with ustult ultisols. These associations suggest that soil texture, mineral composition, drainage, and microclimatic buffering substantially influence fungal persistence and reproduction. Experimental studies have shown that *P. brasiliensis* grows successfully in both sandy and clay-rich soils with elevated moisture and high exchangeable aluminium content (Terçarioli et al., 2007), while *P. lutzii* exhibits comparatively reduced mycelial growth and conidial production under soil-mimicking conditions, consistent with a more restricted saprobic distribution (Hrycyk et al., 2018). Molecular detection studies further confirm frequent occurrence in soils, particularly within armadillo burrows, supporting a geophilous lifestyle (Arantes et al., 2016). Together, these findings highlight the joint role of macroclimatic and fine-scale edaphic variables and underscore the value of detailed soil surveys (e.g., pH, aluminium content, organic matter, porosity) for future modelling efforts.

Ordination analyses (PCoA) indicate that ecological differentiation among lineages is best interpreted as a continuum with partial structuring rather than discrete clusters. While substantial overlap occurs in environmental space, lineage tendencies remain detectable: S1 occupies the broadest ecological space, PS3 overlaps strongly with S1 but within a reduced niche domain, PS2 is comparatively more restricted and differentiated, and *P. lutzii* forms a partially overlapping but distinct environmental grouping. Hierarchical clustering further supports this intermediate structure, recovering a pattern broadly consistent with (i) S1–PS3 affinity, (ii) PS2 differentiation, and (iii) *P. lutzii* intermediate placement between major *P. brasiliensis* lineages. This structure aligns with phylogenetic and population genetic evidence indicating that *P. brasiliensis* comprises multiple cryptic lineages (S1, PS2, PS3) (Matute et al., 2006), with deep genomic divergence and limited gene flow even in sympatric regions (Mavengere et al., 2020). Nuclear multilocus analyses consistently support these as distinct clades, with PS2 branching earliest and S1 forming a sister group to the PS3/PS4 cluster, reinforcing long-term evolutionary separation.

Importantly, these ecological patterns likely reflect both evolutionary history and phenotypic/ecophysiological traits, including differences in yeast morphology, sporulation capacity, and growth performance (Roberto et al., 2021; Siqueira et al., 2015). However, ordination results are inherently constrained by the environmental variables included and may not fully capture microhabitat variation, host–soil interactions, or spatial sampling bias.

Climatic gradient analyses show that ecological structure is shaped by interacting temperature, precipitation, and aridity systems. Temperature-related variables (bio11, bio6, bio1) primarily define the main environmental axis, whereas precipitation variables (bio12–bio18)

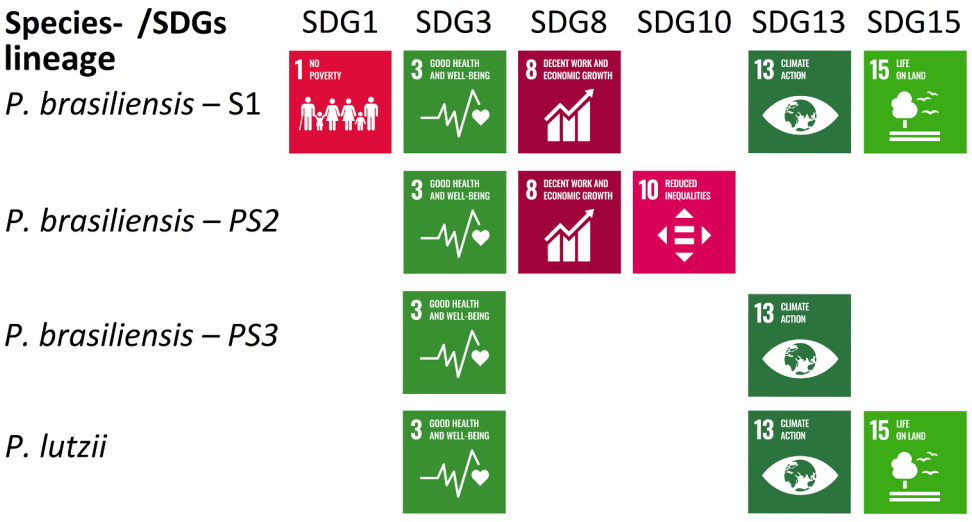

Fig. 9. SDG-oriented conclusions by *Paracoccidioides* species and phylogenetic lineages.

contribute more strongly to secondary gradients. The Köppen Aridity Index (KAI) connects these dimensions, reflecting coupled thermal and hydrological constraints. Machine-learning results are consistent with this framework, identifying temperature, elevation, and soil type as the main organising factors, while precipitation plays a secondary modulatory role. Decision-tree models consistently highlight bio11, elevation, and udult ultisols as key predictors, suggesting that topography and temperature define fundamental niche limits, while soil properties and precipitation refine realised distributions.

Ensemble models provide further insight into the degree of ecological separability among lineages. Species-level comparisons between *P. brasiliensis* and *P. lutzii* show strong predictive separability and high model agreement, indicating clear ecological differentiation at the species level. In contrast, lineage-level comparisons, particularly those involving PS2, show greater variability among models and reduced consistency, suggesting partial overlap in environmental space and more gradual ecological differentiation rather than discrete partitioning.

This interpretation is also reflected in elevated RMSE values observed in PS2–*P. lutzii* comparisons, which likely indicate instability in class boundaries rather than true predictive error. The coexistence of high classification accuracy and relatively higher RMSE further supports continuous ecological structuring rather than sharply defined groups. Exceptionally high accuracies in some ensemble models (e.g. Extra Trees >0.99) should be interpreted cautiously, as they may reflect limitations related to sample size, spatial autocorrelation, and predictor collinearity. Overall, the models capture relative environmental discriminability rather than strict ecological boundaries.

The broad ecological tolerance of S1, spanning elevations, climate zones, and soil types, likely facilitates its wider geographic distribution and potentially higher human exposure risk (Terçarioli et al., 2007; Hahn et al., 2019). In contrast, PS2 and *P. lutzii* remain associated with more environmentally constrained regions characterized by specific climate and edaphic conditions consistent with field and molecular data (Hrycyk et al., 2018; Mavengere et al., 2020). Nevertheless, the observed overlap among lineages indicates that ecological differences should be interpreted as gradual positioning along shared gradients rather than strict niche isolation. Overall, these patterns inform targeted epidemiological surveillance, risk mapping, and public health strategies. Exceptions, such as detections of *P. lutzii* outside predicted niches, highlight the role of microhabitat variability, host ecology, and anthropogenic change, emphasizing the need for mechanistic studies, microclimatic sampling, and temporally explicit modelling.

Climatic suitability patterns indicate a broad potential distribution of *P. brasiliensis* across warm and humid regions of central and Southeastern South America, including Brazil, northern Argentina, and Paraguay, consistent with epidemiological observations (Terçarioli et al., 2007; Hahn et al., 2019). The Brazilian Highlands and portions of the Amazon Basin appear particularly suitable, reflecting ecological plasticity. Soil and armadillo-mediated dispersal may further facilitate its wide distribution (Mavengere et al., 2020). Within the complex, S1 retains the widest suitability, whereas PS2 shows a stronger association with cooler, stable, and humid Atlantic Forest environments. In contrast, *P. lutzii* remains largely restricted to central-western Brazil, consistent with its narrower ecological niche.

The spatial heterogeneity among lineages carries important epidemiological implications. Regions with high environmental suitability may correspond to elevated exposure risk, particularly under land-use change, deforestation, agriculture, or soil disturbance. Conversely, climatic constraints may limit expansion of PS2 and *P. lutzii*, contributing to observed geographic structuring and genomic divergence. These patterns underscore the importance of integrating ecological gradients into disease surveillance and risk mapping frameworks.

The studied xenarthrans exhibited substantial environmental overlap with *Paracoccidioides* lineages, rather than being highly distinct. *D. novemcinctus* occupied a broad and heterogeneous environmental space, showing high overlap with all fungal lineages (D = 0.91–0.97; H = 0.036–0.091), consistent with its wide geographic range and ecological flexibility (Feng and Papeş, 2015). Overall, *D. novemcinctus* occupies environmental niches that are extensive and largely shared with multiple *Paracoccidioides* lineages, supporting its role as a major reservoir host and suggesting that host ecological breadth facilitates pathogen persistence across diverse habitats (Bagagli et al., 1998, 2003; Cocio et al., 2020). More broadly, host environmental niches appear to overlap substantially with pathogen distributions, reinforcing the ecological coupling between reservoir hosts and environmental fungi across diverse habitats.

A key limitation of this study relates to the incomplete and spatially biased nature of occurrence data. Because PCM is underdiagnosed and underreported, available records likely underestimate the true distribution of *Paracoccidioides* spp. Consequently, model outputs should be interpreted as environmental suitability estimates rather than definitive presence–absence patterns. Differences between predicted suitability and reported clinical cases likely reflect sampling bias, reporting gaps, and georeferencing limitations rather than true absence of the fungus.

5. Conclusion

Integrative analyses indicate that the distribution and ecological differentiation of *Paracoccidioides* lineages are shaped by interacting effects of elevation, climate, soil characteristics, and host associations. Rather than forming strictly discrete ecological units, these factors generate a dynamic environmental continuum structured by multidimensional gradients, within which line-age-level ecological differentiation emerges. Within this framework, S1 exhibits the broadest ecological tolerance across elevations, climates, and soil types, supporting its wide geographic distribution and potentially higher human exposure risk. In contrast, PS2, PS3 and *P. lutzii* occupy comparatively more restricted and environmentally specific niches, particularly associated with mid-elevation and edaphically constrained conditions, reflecting varying degrees of ecological specialisation. Precipitation and soil variables emerge as key determinants of realised niches, highlighting the importance of seasonal rainfall patterns, aridity balance, and soil composition in sustaining environmental persistence of *Paracoccidioides* populations. Quantitative niche-overlap metrics (Schoener's D and Hellinger distance) further demonstrate substantial but incomplete overlap among lineages and with xenarthran hosts, particularly *D. novemcinctus*, providing a robust framework for evaluating pathogen–host co-occurrence and environmental exposure patterns. These results are consistent with the role of reservoir host ecology in maintaining fungal persistence across heterogeneous landscapes.

Funding

This paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences under grant number bo/00896/24/5, and by the RRF-2.3.1-21-2022-00014 National Multidisciplinary Laboratory for Climate Change project.

CRediT authorship contribution statement

Attila J. Trájer: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Katalin Vehovszky:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.actatropica.2026.108209.

Data availability

Data will be made available on request.

References

Almeida, F., 1930. Estudos comparativos do granuloma coccidióidico nos Estados Unidos e no Brasil. Novo gênero para o parasito brasileiro. [Comparative studies of coccidioides granuloma in the United States and Brazil. A new genus for the Brazilian parasite.]. An. Fac. Med. Univ. São Paulo 5, 125–141. In Portugal.

Andrade-Silva, J., Andrade-Silva, L.E., Paes, H.C., Alves, L., Rosa, A., Tenório, B.G., Ferreira, M.S., Felipe, M.S.S., Teixeira, M.D.M., Silva-Vergara, M.L., 2021. Molecular epidemiology of *Paracoccidiodes* spp. recovered from patients with paracoccidioidomycosis in a teaching hospital from Minas Gerais State of Brazil. PLoS Negl. Trop. Dis. 15 (11), e0009956. https://doi.org/10.1371/journal.pntd.0009956.

Arantes, T.D., Theodoro, R.C., Teixeira, M.D.M., Bosco, S.D.M.G., Bagagli, E., 2016. Environmental mapping of *Paracoccidioides* spp. in Brazil reveals new clues into genetic diversity, biogeography and wild host association. PLoS Negl. Trop. Dis. 10, e0004606. https://doi.org/10.1371/journal.pntd.0004606.

Bagagli, E., Sano, A., Coelho, K.I., Alquati, S., Miyaji, M., De Camargo, Z.P., Gomes, G.M., Franco, M., Montenegro, M.R., 1998. Isolation of *Paracoccidioides brasiliensis* from armadillos (*Dasypus noveminctus*) captured in an endemic area of paracoccidioidomycosis. Am. J. Trop. Med. Hyg. 58, 505–512.

Bagagli, E., Franco, M., Bosco, S.D.M., Hebeler-Barbosa, F., Trinca, L.A., Montenegro, M. R., 2003. High frequency of *Paracoccidioides brasiliensis* infection in armadillos (*Dasypus novemcinctus*): an ecological study. Med. Mycol. 41, 217–223. https://doi.org/10.1080/1369378031000159736 8.

Bagagli, E., Theodoro, R.C., Bosco, S.M.G., McEwen, J.G., 2008. *Paracoccidioides brasiliensis*: phylogenetic and ecological aspects. Mycopathologia 165, 197–207. https://doi.org/10.1007/s11046-007-9050-7.

Bagagli, E., Matute, D.R., Garces, H.G., Tenorio, B.G., Garces, A.G., Alves, L.G.D.B., Yamauchi, D.H., Hrycyk, M.F., Barker, B.M., Teixeira, M.D.M., 2021. *Paracoccidioides brasiliensis* isolated from nine-banded armadillos (*Dasypus novemcinctus*) reveal population structure and admixture in the Amazon Basin. J. Fungi 7, 54. https://doi.org/10.1007/s11046-007-9050-7.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5, 1–12. https://doi.org/10.1038/sdata.2018.214.

Benedict, K., Richardson, M., Vallabhaneni, S., Jackson, B.R., Chiller, T., 2017. Emerging issues, challenges, and changing epidemiology of fungal disease outbreaks. Lancet Infect. Dis. 17, e403–e411. https://doi.org/10.1016/S1473-3099(17)30443-7.

Bosco, S.D.M.G., Bagagli, E., 2018. Paracoccidioidomycosis in animals and humans. In: Seyedmousavi, S., de Hoog, G., Guillot, J., Verweij, P. (Eds.), Emerging and Epizootic Fungal Infections in Animals. Springer, Cham. https://doi.org/10.1007/978-3-319-72093-7₆.

Calle, D., Rosero, D.S., Orozco, L.C., Camargo, D., Castañeda, E., Restrepo, A., 2001. Paracoccidioidomycosis in Colombia: an ecological study. Epidemiol. Infect. 126 (2), 309–315. https://doi.org/10.1017/S0950268801005052.

Cerasoli, F., D'Alessandro, P., Biondi, M., 2022. Worldclim 2.1 versus Worldclim 1.4: climatic niche and grid resolution affect between-version mismatches in habitat suitability models predictions across Europe. Ecol. Evol. 12, e8430. https://doi.org/10.1002/ece3.8430.

Cocio, T.A., Nascimento, E., Kress, M.R., Bagagli, E., Martinez, R., 2020. Characterization of a *Paracoccidioides* spp. strain from southeastern Brazil genotyped as *Paracoccidioides restrepiensis* (PS3) and review of this phylogenetic species. Genet. Mol. Biol. 43, e20190201. https://doi.org/10.1590/1678-4685-GMB-2019-0201.

De Macedo, P.M., Teixeira, M.D.M., Barker, B.M., Zancope-Oliveira, R.M., Almeida-Paes, R., Francesconi do Valle, A.C., 2019. Clinical features and genetic background of the sympatric species *Paracoccidioides brasiliensis* and *Paracoccidioides americana*. PLoS Neglect. Trop. Dis. 13, e0007309. https://doi.org/10.1371/journal.pntd.0007309.

Denning, D.W., 2024. Global incidence and mortality of severe fungal disease. Lancet Infect. Dis. 24, e428–e438. https://doi.org/10.1016/S1473-3099(23)00692-8.

Fabris, L.R., de Oliveira, N.G., Bortolomai, B.E., Batista, L.C.F., Sobral, M.H., Ribeiro, A. A., Andrade, U.V., Filho, A.C.P., Carvalho, L.R.D., Baptista, I.M.F.D., Mendes, R.P., Paniago, A.M.M., 2024. The effect of geoclimatic factors on the distribution of paracoccidioidomycosis in Mato Grosso do Sul, Brazil. J. Fungi 10, 165. https://doi.org/10.3390/jof10030165.

Feng, X., Papeş, M., 2015. Ecological niche modelling confirms potential north-east range expansion of the nine-banded armadillo (*Dasypus novemcinctus*) in the USA. J. Biogeogr. 42, 803–807. https://doi.org/10.1111/jbi.12427.

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37, 4302–4315. https://doi.org/10.1002/joc.5086.

Gottdenker, N.L., Streicker, D.G., Faust, C.L., Carroll, C.R., 2014. Anthropogenic land use change and infectious diseases: a review of the evidence. EcoHealth 11, 619–632. https://doi.org/10.1007/s10393-014-0941-z.

Gower, J.C., 2014. Principal coordinates analysis. Wiley StatsRef: statistics reference online, 1–7. https://doi.org/10.1002/9781118445112.stat05670.pub2.

Hahn, R.C., Rodrigues, A.M., Della Terra, P.P., Nery, A.F., Hoffmann-Santos, H.D., Gois, H.M., Fontes, C.J.F., de Camargo, Z.P., 2019. Clinical and epidemiological features of paracoccidioidomycosis due to *Paracoccidioides lutzii*. PLoS Neglect. Trop. Dis. 13, e0007437. https://doi.org/10.1371/journal.pntd.0007437.

Hahn, R.C., Hagen, F., Mendes, R.P., Burger, E., Nery, A.F., Siqueira, N.P., Guevara, A., Rodrigues, A.M., de Camargo, Z.P., 2022. Paracoccidioidomycosis: current status and future trends. Clin. Microbiol. Rev. 35, e0023321. https://doi.org/10.1128/cmr.00233-21.

Hrycyk, M.F., Garcia Garces, H., Bosco, S.D.M.G., de Oliveira, S.L., Marques, S.A., Bagagli, E., 2018. Ecology of *Paracoccidioides brasiliensis, P. lutzii* and related species: infection in armadillos, soil occurrence and mycological aspects. Med. Mycol. 56, 950–962. https://doi.org/10.1093/mmy/myx142.

Liu, C., Newell, G., White, M., Machunter, J., 2025. Improving the estimation of the Boyce index using statistical smoothing methods for evaluating species distribution models with presence-only data. Ecography 2025, e07218. https://doi.org/10.1111/ecog.07218. Digital Object Identifier (DOI).

MacFerrin, M., Amante, C., Carignan, K., Love, M., Lim, E., 2024. The earth topography 2022 (ETOPO 2022) global dem dataset. Earth Syst. Sci. Data Discuss. 2024, 1–24. https://doi.org/10.5194/essd-17-1835-2025.

Martinez, R., 2015. Epidemiology of paracoccidioidomycosis. Rev. Inst. Med. Trop. Sao Paulo 57, 11–20. https://doi.org/10.1590/S0036-46652015000700004.

Mattos, K., Cocio, T.A., Chaves, E.G.A., Borges, C.L., Venturini, J., de Carvalho, L.R., Mendes, R.P., Paniago, A.M., Weber, S.S., 2021. An update on the occurrence of *paracoccidioides* species in the Midwest region, Brazil: molecular epidemiology, clinical aspects and serological profile of patients from Mato Grosso do Sul State. PLoS Neglect. Trop. Dis. 15, e0009317. https://doi.org/10.1371/journal.pntd.0009317.

Matute, D.R., McEwen, J.G., Puccia, R., Montes, B.A., San-Blas, G., Bagagli, E., Rauscher, J.T., Restrepo, A., Morais, F., Niño-Vega, G., Taylor, J.W., 2006. Cryptic speciation and recombination in the fungus *Paracoccidioides brasiliensis* as revealed by gene genealogies. Molecular Biol. Evol. 23, 65–73. https://doi.org/10.1093/molbev/msj008.

Mavengere, H., Mattox, K., Teixeira, M.M., Sepúlveda, V.E., Gomez, O.M., Hernandez, O., McEven, J., Matute, D.R., 2020. *Paracoccidioides* genomes reflect high levels of species divergence and little interspecific gene flow. MBio 11, 10–1128. https://doi.org/10.1128/mbio.01999-20.

Muñoz, J.F., Farrer, R.A., Desjardins, C.A., Gallo, J.E., Sykes, S., Sakthikumar, S., Misas, E., Whiston, E.A., Bagagli, E., Soares, C.M.A., Teixeira, M.D.M., Taylor, J.W., Clay, O.K., McEwen, J.G., Cuomo, C.A., 2016. Genome diversity, recombination, and virulence across the major lineages of *Paracoccidioides*. Msphere 1, 10–1128. https://doi.org/10.1128/msphere.00213-16.

Panagos, P., Van Liedekerke, M., Borrelli, P., Köninger, J., Ballabio, C., Orgiazzi, A., Lugato, E., Liakos, L., Hervas, J., Jones, A., Montanarella, L., 2022. European soil data centre 2.0: soil data and knowledge in support of the EU policies. Eur. J. Soil Sci. 73, e13315. https://doi.org/10.1111/ejss.13315.

Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J.H., Carioli, A., Corbane, C., Dijkstra, L., Florio, P., Friedrich, H.K., Gao, J., Leyk, S., Lu, L., Maffenini, L., Mari-Rivero, I., Melchiorri, M., Syrris, V., Van Den Hoek, J., Kemper, T., 2024. Advances on the global human settlement layer by joint assessment of earth observation and population survey data. Int. J. Digit. Earth 17, 2390454. URL. https://developers.google.com/earth-engine/datasets/catalog/JRC_GHSL_P2023A_GHS_SMOD_V2-0?utm_source=chatgpt.com.

Pinheiro, B.G., Hahn, R.C., Camargo, Z.P.D., Rodrigues, A.M., 2020. Molecular tools for detection and identification of *Paracoccidioides* species: current status and future perspectives. J. Fungi 6, 293. https://doi.org/10.3390/jof6040293.

Quan, C., Han, S., Utescher, T., Zhang, C., Liu, Y.S.C., 2013. Validation of temperature–precipitation based aridity index: paleoclimatic implications. Palaeogeogr. Palaeoclimatol. Palaeoecol. 386, 86–95. https://doi.org/10.1016/j.palaeo.2013.05.008.

Restrepo-Moreno, A., 2018. Ecology of *Paracoccidioides brasiliensis*. Paracoccidioidomycosis. CRC Press, pp. 121–130.

Richini-Pereira, V.B., Bosco, S.M., Theodoro, R.C., Lígia Barrozo, L., Pedrini, S.C.B., Rosa, P.S., Bagagli, E., 2009. Importance of xenarthrans in the eco-epidemiology of *Paracoccidioides brasiliensis*. BMC Res. Notes 2, 228. https://doi.org/10.1186/1756-0500-2-228.

Roberto, T.N., De Carvalho, J.A., Beale, M.A., Hagen, F., Fisher, M.C., Hahn, R.C., Camargo, Z.P.D., Rodrigues, A.M., 2021. Exploring genetic diversity, population structure, and phylogeography in *Paracoccidioides* species using AFLP markers. Stud. Mycol. 100, 100131. https://doi.org/10.1016/j.simyco.2021.100131.

Rodrigues, M.L., Nosanchuk, J.D., 2021. Fungal diseases as neglected pathogens: a wake-up call to public health officials. Advances in Clinical Immunology, Medical Microbiology, COVID-19, and Big Data. Jenny Stanford Publishing, pp. 399–411.

Rokas, A., 2022. Evolution of the human pathogenic lifestyle in fungi. Nat. Microbiol. 7, 607–619. https://doi.org/10.1038/s41564-022-01112-0.

San-Blas, G., Niño-Vega, G., Iturriaga, T., 2002. *Paracoccidioides brasiliensis* and paracoccidioidomycosis: molecular approaches to morphogenesis, diagnosis, epidemiology, taxonomy and genetics. Med. Mycol. 40, 225–242. https://doi.org/10.1080/mmy.40.3.225.242.

Saputra, D.M., Saputra, D., Oswari, L.D., 2020. Effect of distance metrics in determining k-value in k-means clustering using elbow and silhouette method. In: Sriwijaya international conference on information technology and its applications (SICONIAN 2019). Atlantis Press, p. 341.

Sharpton, T.J., Stajich, J.E., Rounsley, S.D., Gardner, M.J., Wortman, J.R., Jordar, V.S., Maiti, R., Kodira, C.D., Neafsey, D.E., Zeng, O., Hung, C.-Y., McMahan, C.,

Muszewska, A., Grynberg, M., Mandel, M.A., Kellner, E.M., Barker, B.M., Galgiani, J. N., Orbach, M.J., Kirkland, T.N., Cole, G.T., Henn, M.R., Birren, B.W., Taylor, J.W., 2009. Comparative genomic analyses of the human fungal pathogens coccidioides and their relatives. Genome Res. 19, 1722–1731. https://doi.org/10.1101/gr.087551.108.

Siqueira, I.M., Fraga, C.L.F., Amaral, A.C., Souza, A.C.O., Jerônimo, M.S., Correa, J.R., Souza, A.C.O., Jerônimo, M.S., Correa, J.R., Magalhães, K.G., Inácio, C.A., Ribeiro, A.M., Burguel, P.H., Felipe, M.S., Tavares, A.H., Bocca, A.L., 2015. Distinct patterns of yeast cell morphology and host responses induced by representative strains of *Paracoccidioides brasiliensis* (Pb18) and *Paracoccidioides lutzii* (Pb01). Sabouraudia 54, 177–188. https://doi.org/10.1093/mmy/myv072.

Teixeira, M.D.M., Theodoro, R.C., Oliveira, F.F.M.D., Machado, G.C., Hahn, R.C., Bagagli, E., San-Blas, G., Felipe, M.S.S., Soares Felipe, M.S., 2014. *Paracoccidioides lutzii* sp. nov.: biological and clinical implications. Med. Mycol. 52, 19–28. https://doi.org/10.3109/13693786.2013.794311.

Teixeira, M.D.M., Cattana, M.E., Matute, D.R., Muñoz, J.F., Arechavala, A., Isbell, K., Schipper, R., Santiso, G., Tracogna, F., Sosa, M.D.L.Á., Cech, N., Alvarado, P., Barreto, L., Chacón, Y., Ortellado, J., Lima, C.M.D., Chang, M.R., Niño-Vega, G., Yasuda, M.A.S., Felipe, M.S.S., Giusiano, G., 2020. Genomic diversity of the human pathogen *Paracoccidioides* across the South American continent. Fungal Genet. Biol. 140, 103395. https://doi.org/10.1016/j.fgb.2020.103395.

Terçarioli, G.R., Bagagli, E., Reis, G.M., Theodoro, R.C., Bosco, S.D.M.G., Macoris, S.A.D. G., Richini-Pereira, V.B., 2007. Ecological study of *Paracoccidioides brasiliensis* in soil: growth ability, conidia production and molecular detection. BMC Microbiol. 7, 92. https://doi.org/10.1186/1471-2180-7-92.

Theodoro, R.C., Teixeira, M.D.M., Felipe, M.S.S., Paduan, K.D.S., Ribolla, P.M., San-Blas, G., Bagagli, E., 2012. Genus *Paracoccidioides*: species recognition and biogeographic aspects. PloS One 7, e37694. https://doi.org/10.1371/journal.pone.0037694.

Turissini, D.A., Gomez, O.M., Teixeira, M.M., McEwen, J.G., Matute, D.R., 2017. Species boundaries in the human pathogen Paracoccidioides. Fungal Genet. Biol. 106, 9–25. https://doi.org/10.1016/j.fgb.2017.05.00.

Williams, S.L., Toda, M., Chiller, T., Brunkard, J.M., Litvintseva, A.P., 2024. Effects of climate change on fungal infections. PLoS Pathog. 20, e1012219. https://doi.org/10.1371/journal.ppat.1012219.

Zhang, L., Wang, L., 2023. Optimization of site investigation program for reliability assessment of undrained slope using Spearman rank correlation coefficient. Comput. Geotech. 155, 105208. https://doi.org/10.1016/j.compgeo.2022.105208.

Zhou, J., Zhang, X., Qu, Z., Zhang, C., Wang, F., Gao, T., Yao, Y., Liang, J., 2024. Progress in research on prevention and control of crop fungal diseases in the context of climate change. Agriculture 14, 1108. https://doi.org/10.3390/agriculture14071108.