UV index and climate seasonality explain fungal community turnover in global drylands

Eleonora Egidi1,2 | Manuel Delgado-Baquerizo3,4 | Miguel Berdugo5,6 | Emilio Guirado5 | Davide Albanese7 | Brajesh K. Singh1,2 | Claudia Coleine8

1Hawkesbury Institute for the Environment, Western Sydney University, Penrith, New South Wales, Australia
2Global Centre for Land-Based Innovation, Western Sydney University, Penrith, New South Wales, Australia
3Laboratorio de Biodiversidad y Funcionamiento Ecosistémico, Instituto de Recursos Naturales y Agrobiología de Sevilla (IRNAS), CSIC, Seville, Spain
4Unidad Asociada CSIC-UPO (BioFun), Universidad Pablo de Olavide, Seville, Spain
5Instituto Multidisciplinar para el Estudio del Medio "Ramon Margalef", Universidad de Alicante, Edificio Nuevos Institutos, San Vicente del Raspeig, Spain
6Institut Department of Environmental Systems Science, ETH Zürich, Zurich, Switzerland
7Research and Innovation Centre, Fondazione Edmund Mach (FEM), San Michele all’Adige, Italy
8Department of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy

Correspondence
Eleonora Egidi, Hawkesbury Institute for the Environment, Western Sydney University, Penrith, NSW, Australia. Email: e.egidi@westernsydney.edu.au
Claudia Coleine, Department of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy. Email: coleine@unitus.it

Handling Editor: Janne Soininen

Abstract
Aim: Fungi are major drivers of ecosystem functioning. Increases in aridity are known to negatively impact fungal community composition in dryland ecosystems globally; yet, much less is known on the potential influence of other environmental drivers, and whether these relationships are linear or nonlinear.
Location: Global.
Major taxa studied: Fungi.
Methods: We re-analysed multiple datasets from different dryland biogeographical regions, for a total of 912 samples and 1,483 taxa. We examined geographical patterns in community diversity and composition, and spatial, edaphic and climatic factors driving them.
Results: UV index, climate seasonality, and sand content were the most important environmental predictors of community shifts, showing the strongest association with the richness of putative plant pathogens and saprobes. Important nonlinear relationships existed with each of these fungal guilds, with increases in UV and temperature seasonality above 7.5 and 900 SD (standard deviation x 100 of the mean monthly temperature), respectively, being associated with an increased probability of plant pathogen and unspecified saprotroph occurrence. Conversely, these environmental parameters had a negative relationship with litter and soil saprotroph richness. Consequently, these ecological groups might be particularly sensitive to shifts in UV radiation and climate seasonality, which is likely to disturb current plant–soil dynamics in drylands.
Main conclusions: Our synthesis integrates fungal community data from drylands across the globe, allowing the investigation of fungal distribution and providing the first evidence of shifts in fungal diversity and composition of key fungal ecological groups along diverse spatial, climatic and edaphic gradients in these widely distributed ecosystems. Our findings imply that shifts in soil structure and seasonal climatic patterns induced by global change will have disproportionate consequences for the
1 | INTRODUCTION

Drylands are the largest terrestrial biome (covering about 41% of the land surface and supporting 40 and 35% of the global population and global diversity, respectively) and are expected to expand further up to 56% by the end of the century (Cherlet et al., 2018). Drylands play key roles in regulating the global carbon, nitrogen and water cycles, and are thus fundamental for sustaining life on Earth (Maestre et al., 2021). Due to their extreme temperatures, low and variable rainfall, and low soil fertility, drylands are particularly sensitive to changes in climate that lead to increased aridity (i.e., precipitation/potential evapotranspiration; Berdugo et al., 2020).

Fungi are paramount components of and drive critical ecosystem services in drylands, contributing to the formation of fertile islands (Cherlet et al., 2018), nutrient cycling and climate regulation (Delgado-Baquerizo et al., 2017), with a major role in dryland primary production (Rudgers et al., 2018) and pedogenesis (Coleine et al., 2021). Key fungal groups include pathogens, mutualistic symbionts of both plants and animals, lichenized fungi, as well as soil and litter saprobes.

Patterns of fungal diversity and community turnover are particularly little explored in drylands. Most previous studies on the biogeography and ecological attributes of fungal communities in dry ecosystems on a global scale have focussed on their relationships with aridity, given its role as a key driver of dryland ecology (Berdugo et al., 2020; Egidi, Delgado-Baquerizo, et al., 2019; Maestre et al., 2015). However, other climatic and environmental factors are potentially important in predicting fungal diversity and distributions in global drylands, especially at the level of guilds. Drylands are characterized by extreme temperatures, low and variable rainfall, and low soil fertility, which can interact to drive multiple ecosystem attributes. For example, temperature and precipitation seasonality regulate plant cover dynamics and productivity in arid systems (Palmquist et al., 2021), which in turn can influence soil physical attributes, such as soil moisture, pH, structure or carbon content, important for the distribution of saprotrophs, pathogens, and symbiotrophs (Egidi, Wood, et al., 2019; Feng et al., 2022). Consequently, community turnover driven by seasonal processes, rather than aridity alone, may be particularly important for fungi in these systems (Berdugo et al., 2022). Similarly, solar UV radiation is a primary abiotic driver of litter and soil organic carbon decomposition in many arid and semi-arid ecosystems (Bornman et al., 2015; Zepp et al., 2007), suggesting a potential major influence upon the occurrence of decomposers and plant-associated fungi (Paul & Gwynn-Jones, 2003).

Examining the major patterns in fungal community shifts is important to elucidate the effects of environmental change for community composition or diversity. In particular, the climate hypothesis (O’Brien, 1998) indicates that species richness and composition over large scales are linked to climate, whereby species richness and community turnover along wide gradients are primarily controlled by the availability of water (e.g., precipitation and evapotranspiration) and ambient energy (e.g., temperature) (Hawkins et al., 2003). Given the ecological and economic significance of drylands, and the global role of fungi in regulating their functions, it is critical to identify the climatic and environmental factors associated with distributions of fungal communities, and most importantly, to test whether the dependence of fungi on those drivers is linear or nonlinear. The latter is important because nonlinear associations between fungal distributions and environmental predictors may signal particular environmental scenarios of exacerbated sensitivity to environmental pressures, with potential implications for ecosystem functioning (Pausas & Bond, 2020). A better understanding of the drivers shaping the biogeography of dryland soil fungi can thus improve our ability to predict their fate under global change, and therefore inform future conservation and management policies. Towards this aim, we re-analysed multiple datasets from different dryland biogeo graphical regions, merging sequencing data from a wide range of ecosystems and climates (i.e. hot, temperate and cold drylands) to encompass a representative plethora of all dryland sub-types (i.e. from hyperarid to dry sub-humid). We generated a database of 1,473 fungal genera from a total of 912 individual topsoil samples (top 0 to 5–15 cm) from all continents, including Antarctica. We examined geographical patterns in fungal assemblages and the main environmental (spatial, edaphic and climatic) factors driving them in order to establish where important changes in community composition occur along this range of environmental gradients. We expect that shifts in water and energy availability related to climatic variables will help in identifying predictors of fungal diversity and distribution of key ecological groups in drylands on a global scale.

2 | MATERIALS AND METHODS

2.1 | Literature selection

We created a dataset by collecting published literature on studies of soil fungal communities in drylands across the globe. For this literature search, we conducted searches in the Web of Knowledge (https://www.webofknowledge.com), Google Scholar
Appendix S1. It should be noted the studies we considered included request). A list of the data sources is found in Supporting Information data (the data were neither public nor provided by the authors upon some of the publications, we were not able to obtain sequencing Information Figure S1, and Table S1 for more details); however, for according to the Environment Ontology (https://www.ontobee.org/ assigned to samples; and (f) the samples could be assigned to biomes worldwide* OR hot* OR cold*) AND (drylands* OR deserts* OR arid/semi-arid/hyperarid/dry sub-humid region/area*) (soil* OR biocrusts* OR "biological crusts") AND (fungi* OR fung* OR ITS* OR ITS1* OR ITS2* OR fungal communities*) AND (high-throughput sequence* OR amplicon sequencing* or metabarcoding* OR Illumina sequenc*). Additionally, we included studies found in the references of the papers returned by the database. A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (Moher et al., 2009; Supporting Information Figure S1) is provided to summarize the literature search and selection processes. Briefly, we explored 116 papers. The following selection criteria were used for the inclusion of samples (and, consequently, studies) in the dataset: (a) samples came from soils or biocrusts; (b) the precise geographical location of each sample was recorded (GPS coordinates) or provided by the authors of the study upon request; (c) the whole fungal community was subject to amplicon sequencing (studies using group specific primers were excluded); (d) the internal transcribed spacer regions (ITS1, ITS2 or both) were amplified and sequenced on Illumina platforms; (e) sequencing data (either in fasta or fastq format) were publicly available or provided by the authors of the study upon request, and the sequences were unambiguously assigned to samples; and (f) the samples could be assigned to biomes according to the Environment Ontology (https://www.ontobee.org/ontology/ENVO). In total, 13 sequencing studies contained samples that matched our criteria (see PRISMA flow diagram in Supporting Information Figure S1, and Table S1 for more details); however, for some of the publications, we were not able to obtain sequencing data (the data were neither public nor provided by the authors upon request). A list of the data sources is found in Supporting Information Appendix S1. It should be noted the studies we considered included samples collected during different sampling campaigns but information on sampling year was not available for individual samples for most locations. However, seasonal variability is now known to be far less important than environmental variability both at the local and global scale (Carini et al., 2020); consequently, sampling year should have a negligible effect when considering sites spanning many regions of the globe. Bioinformatics details are provided in Supporting Information Note S1.

2.2 Environmental variable selection

We performed a comprehensive meta-study of data published on the composition of soil fungal communities in drylands across the world. This approach enabled us to re-analyse multiple datasets from different biogeographical regions and biomes and compile a large dataset of fungal taxa distribution worldwide. In total, 13 studies, encompassing over 912 top-soil (0 to 5–10 cm depth) sampling points, were identified and included in the analysis; this allowed us to encompass all continents (including Antarctica; Supporting Information Figure S1), spanning a wide range of environmental conditions. The final sample list included all dryland subtypes (hyperarid, aridity index (AI). 0–.05, n = 42; arid, AI = .05–.20, n = 274; semi-arid, AI = .20–.50; n = 336; dry subhumid, AI = .50–.65; n = 264).

Metadata were either collected from the published papers and/or public repositories where they were submitted by the authors, or in a few cases from the authors of individual studies upon request and are included in Supporting Information (Table S1). Additional metadata were collected from the WorldClim database ([https://www.worldclim.org]; –1 km resolution (Fick & Hijmans, 2017), similar to Sanderman et al., 2017; Van Den Hoogen et al., 2019), and included spatial, climatic and edaphic parameters. Climatic data included a range of variables related to temperature and precipitation variability that are considered important drivers of fungal distribution at large scales (Vétovsy et al., 2019) – that is, mean annual temperature (MAT), precipitation seasonality (PSEA), temperature seasonality (TSEA), AI (precipitation/potential evapotranspiration) and standardized precipitation-evapotranspiration index (SPEI). The AI was obtained from the global maps of Zomer et al. (2008), which provides the averaged AI of the period 1970–2000, and has a spatial resolution of 30 arc-seconds. We also collected data on the AI from the Global Potential Evapotranspiration database (Trabucco & Zomer, 2019), which is based on interpolations provided by WorldClim. We used AI instead of mean annual precipitation in our study because aridity includes both mean annual precipitation and potential evapotranspiration, and is therefore a more accurate metric of the long-term water availability at each site; moreover, AI is the one used for categorizing drylands and in global reports about desertification and climate change. The SPEI is a multiscale drought index based on climatic data that quantifies temporal variations in water balance and classifies the onset, magnitude and duration of drought conditions relative to regular conditions at a given location (Beguería et al., 2014). SPEI is based on climatic data of monthly precipitation and potential evapotranspiration from the Climatic Research Unit (CRU) TS3.10.01 dataset (https://badc.nerc.ac.uk/; Luo et al., 2016) with FAO-56 Penman-Monteith equation estimation (Harris et al., 2014) at 0.5° spatial resolution. SPEI was used in a monthly time window and averaged for the period 2010–2020 for each location studied (Beguería et al., 2014; Harris et al., 2014; Luo et al., 2016). UV radiation index (UV) was further included given its importance in driving biogeochemical processes in dryland soils (Throop & Archer, 2008; Trabucco & Zomer, 2019). Three important edaphic determinants of fungal biogeography [i.e., % of sand, soil organic carbon (SOC) and pH], obtained from the SoilGrids v2 database, were also included, allowing us to evaluate the importance of soil physico-chemical attributes for fungal distribution in drylands. The global independence of the soil and climatic variables was verified with a Mantel test (999 permutations), which returned a non-significant (p = 1, R = .07 for soil data; p = .65, R = .06 for climatic
The turnover function is measured in dimensionless $R$ units where changes in community compositional turnover or ‘splits’ along predictor gradients. To explore the environmental drivers of fungal community composition across multiple environmental gradients, at the level of individual guilds as well as the mycobiome as a whole when these individual curves are averaged to obtain a global $R^2$ value. Detailed descriptions of these methods can be found in Baker and Hollowed (2014), Ellis et al (2012) and Stephenson et al (2018). Following the GF approach described above, model performance was assessed by the overall goodness-of-fit (R-squared) of predicted against observed values and by the cross-validated out-of-bag $R^2$ values per ecological group, while the significance of each environmental variable was assessed by the relative importance weighted by $R^2$ values (Ellis et al., 2012).

To verify whether results were affected by differing sampling techniques among studies, the model was re-run including soil depth as an additional covariate. Subsequently, to identify critical values along environmental gradients that correspond to changes in composition at the guild level, we plotted their cumulative importance, whereby the shape of the resulting distribution curves describes the magnitude of compositional change along the most important gradients, with a standardized ratio of split density $>1$ indicating the likely presence of community shift (Roland Pitcher et al., 2012). Finally, to further illustrate the directionality of shifts in richness for each individual guild in response to environmental predictors, we ran GradientBoost (GB) models with Shapley additive explanations (SHAP) dependence plots. GB models were run individually for each guild and were done solely for the most important environmental variables and those guilds best explained by the GF models. The SHAP method is derived from game theory and measures how much each feature of a model contributes to the increase or decrease of the probability of a single output with respect to the average of the ones used to train the model (i.e., the richness of a particular ecological group in this case). By plotting the values of predictor versus the associated SHAP values we obtain a response curve analogous to the effects of that predictor over the response variable (i.e., a partial dependence plot). SHAP values are widely used in machine learning, economics, security and ecology (Cha et al., 2021; Foster et al., 2022). SHAP values can be positive or negative, whereby a positive trend indicates that a feature is expected to positively influence the occurrence of a particular guild, and vice versa. Models were built with the ‘xgboost’ package and SHAP values were extracted with the ‘SHAPforxgboost’ package in R.

2.3 | Quantification of ecological turnover and thresholds across environmental gradients

To explore the environmental drivers of distributions of the most common fungal ecological guilds in global drylands, we modelled their occurrence using an approach similar to Chen and Olden (2020). Briefly, we first identified the most common guilds among those occurring in at least 10% of the samples, which resulted in 16 guilds. Rare guilds were excluded because of the uncertainty of environmental prediction for low abundance and frequency taxa. We then explored the most important environmental predictors of fungal ecological turnover by generating a random forest (RF) model fitting a total of 500 trees using the extended modelling procedure available in the R package `gradientForest` (Ellis et al., 2012; Stephenson et al., 2018). Gradient forest (GF) provides a flexible approach to explore multispecies responses between biodiversity and the environment (Ellis et al., 2012). Accounting for complex interactions by fitting multiple regression models, GF quantifies the relative importance of different predictors and their nonlinear correspondence with community compositional turnover or ‘splits’ along predictor gradients. The turnover function is measured in dimensionless $R^2$ units where groups with highly predictive RF models (i.e., high $R^2$ values) have greater influence on the turnover functions than those with lower predictive power (i.e., lower $R^2$). These turnover functions can provide unique insights into the nature of how functional patterns vary along multiple environmental gradients, at the level of individual guilds as well as the mycobiome as a whole when these individual curves are averaged to obtain a global $R^2$ value. Detailed descriptions of these methods can be found in Baker and Hollowed (2014), Ellis et al., (2012) and Stephenson et al. (2018). Following the GF approach described above, model performance was assessed by the overall goodness-of-fit (R-squared) of predicted against observed values and by the cross-validated out-of-bag $R^2$ values per ecological group, while the significance of each environmental variable was assessed by the relative importance weighted by $R^2$ values (Ellis et al., 2012).

To verify whether results were affected by differing sampling techniques among studies, the model was re-run including soil depth as an additional covariate. Subsequently, to identify critical values along environmental gradients that correspond to changes in composition at the guild level, we plotted their cumulative importance, whereby the shape of the resulting distribution curves describes the magnitude of compositional change along the most important gradients, with a standardized ratio of split density $>1$ indicating the likely presence of community shift (Roland Pitcher et al., 2012). Finally, to further illustrate the directionality of shifts in richness for each individual guild in response to environmental predictors, we ran GradientBoost (GB) models with Shapley additive explanations (SHAP) dependence plots. GB models were run individually for each guild and were done solely for the most important environmental variables and those guilds best explained by the GF models. The SHAP method is derived from game theory and measures how much each feature of a model contributes to the increase or decrease of the probability of a single output with respect to the average of the ones used to train the model (i.e., the richness of a particular ecological group in this case). By plotting the values of predictor versus the associated SHAP values we obtain a response curve analogous to the effects of that predictor over the response variable (i.e., a partial dependence plot). SHAP values are widely used in machine learning, economics, security and ecology (Cha et al., 2021; Foster et al., 2022). SHAP values can be positive or negative, whereby a positive trend indicates that a feature is expected to positively influence the occurrence of a particular guild, and vice versa. Models were built with the ‘xgboost’ package and SHAP values were extracted with the ‘SHAPforxgboost’ package in R.

2.4 | Quantification of biodiversity and environmental drivers of fungal community composition

Alpha diversity was estimated using the R ‘iNEXT’ package (Hsieh et al., 2016) and the observed Shannon diversity was used for further statistical analyses. A RF model was built using the ‘randomForest’ package with 500 trees in R to assess the relative contributions of climatic, spatial and edaphic predictors to dryland fungal richness. Statistical analysis was performed to identify how overall richness changed across dryland types by one-way analysis of variance.
(one-way ANOVA) and a pairwise multiple comparison procedure (Wilcoxon test); a small probability p-value (<0.05) indicated a significant difference.

### 2.5 | Mapping the distribution of soil fungal richness across drylands

To predict the extent of the global distribution of soil fungal Shannon diversity, we performed a RF regression analysis (Lahouar & Slama, 2015) using the environmental variables from 743 soil samples for which complete observations were obtained. This excluded most of the Antarctic samples, for which soil data are not available in global databases. This model was developed by finding the set of covariate combinations that most robustly predict the training samples with 999 trees and 999 iterations. To assess the accuracy of the predictions calculated from the RF-based model, we calculated how much the parameter set of predictors differed from the original dataset. We used the Mahalanobis distance of any multidimensional point of the 12 dimensions provided by the exogenous variables to the centre of the known distribution that we previously calculated and the distance of any multidimensional point to the convex hull formed by the soil sample locations used in the model. Subsequently, we used outlier identification to mask our results and provide more reliable predictions in the .9 quantiles of the chi-square distribution with 12 degrees of freedom to which each location belongs (Mallavan et al., 2010). The modelling approach was then validated by returning predicted values (model) versus observed values (soil samples), according to Piñeiro et al. (2008).

### 3 | RESULTS

#### 3.1 | General description of the dataset

Our dataset represents the largest extant fungal community dataset from drylands. Compared to previous large-scale studies focused on fungal diversity in drylands, our survey encompasses all continents, including Antarctica, and spans all dryland subtypes (defined by their aridity ranges), from hyperarid (AI ≤ 0.05, n = 38), to arid (0.05 < AI ≤ 0.2, n = 274), semi-arid (0.2 < AI ≤ 0.5, n = 355) and dry sub-humid (0.5 < AI < 0.65, n = 265) regions of the world. Samples were distributed across cold (n = 378), temperate (n = 458) and hot drylands (n = 71) (Supporting Information Figure S3). The depth of the soil cores varied among studies, and included samples collected at depths of 0–15 cm (n = 273), 0–10 cm (n = 170), 0–7.5 cm (n = 183) and 0–5 cm (n = 180) (Supporting Information Figure S2).

Of the 1,473 genera of fungi retrieved, 60% belonged to Ascomycota, 33% to Basidiomycota, 2.6% to Glomeromycota and 2% to Zygomycota (Supporting Information Figure S5). Out of the 66% (986) of taxa that were assigned to an ecological guild, 34.9% were saprotrophs (including 11.5% wood saprotrophs, 9% litter saprotrophs and 8% soil saprotrophs), 13% plant pathogens, 8% endophytic-mycorrhizal (5% ectomycorrhizal, 1% arbuscular-mycorrhizal and 2% root-rot saprotrophs) and 5% were lichenized (Supporting Information Figure S5). Plant pathogens were mostly dominated by ascomycetes fungi from the classes Dothideomycetes (37.6%) and Leotiomycetes (10%), while ectomycorrhizal fungi and wood, litter and soil saprotrophs were dominated by Agaricomycetes (71, 60.5, 32 and 30%, respectively) (Supporting Information Figure S6).

#### 3.2 | Environmental drivers of functional composition

The relative importance of spatial, edaphic and climatic variables in predicting the composition of the main fungal ecological groups was determined using GF models, which identified the major determinants of community turnover of fungi. The total model prediction performance from the GF analysis (i.e., the proportion of variance explained in a RF) was averaged across the suite of environmental variables from the most common guilds (i.e., among those occurring in at least 10% of the samples), and ranged from 0.01 to 0.12 (R²; Figure 1a). Global community turnover was most strongly associated with the spatial variables (PCNM1 and PCNM2 eigenvector-based vectors, maximum cumulative importance: 0.12 and 0.08, respectively), followed closely by UV index (UV), with a maximum value above 0.07. Importance in relation to other environmental predictors was highest (>0.04) for diurnal temperature range (DTR), sand and temperature seasonality (TSEA), while mean annual temperature (MAT), precipitation seasonality (PSEA), pH, aridity index (AI) and soil organic content (SOC) had the lowest importance values (0.02–0.04) (Figure 1a). An additional model run with soil depth included as an explanatory variable (Supporting Information Figure S7) gave similar results, wherein depth was an important explanatory factor for explaining the compositional shift, but the relative importance of the climatic and edaphic variables did not change substantially.

Then, for each ecological group, we identified the most important predictors of changes in their abundance along spatial, climatic and edaphic gradients. The cumulative model prediction performance of the guilds for which significant predictive power was established (R² > 0) had a range of 0.11–0.86 (R²), with the highest model performance (>0.70) recorded for plant pathogens and soil, litter, wood and unspecified saprotrophs (Figure 1b). The predictive power of PCNM1 and PCNM2 was strongest for these fungi relative to other ecological groups (R² > 0.10; Supporting Information Figure S7). However, plant pathogen and unspecified saprotroph richness were also strongly predicted by UV radiation and sand content (R² values >0.07 and >0.08, respectively), while DTR was the single most important climate predictor of litter saprotroph richness, followed by UV (R² = 0.08 and 0.06, respectively). DTR was also important in predicting soil saprotroph distributions, together with sand content (R² values = 0.06 for both). Conversely, PSEA and TSEA were the strongest predictors of ectomycorrhizal fungi (R² values = 0.04 and 0.05, respectively), with TSEA...
were utilized to identify the occurrence of important community changes at whole community scales along environmental gradients. UV index harboured a major threshold at values $>7$, where most of the data occurred (Figure 2a). This threshold corresponded to a shift in the proportion of most ecological groups, including plant pathogens and litter and dung saprotrophs, as indicated by the steep slope in the cumulative plots (Figure 2b). Along the two second most important climatic variables (DTR and TSEA), we observed multiple subsequent strong splits. The fungal community showed a first response with mean diurnal temperature range $>8 \, ^\circ C$, and then a second shift with mean diurnal range $>14 \, ^\circ C$, the latter mainly corresponding to changes in proportion of a range of saprotrophic fungi (i.e., litter, soil and unspecified saprotrophs). Shifts in lichenized and ectomycorrhizal fungi, as well as animal and plant pathogens and dung saprotrophs, were recorded with a variation of monthly temperature averages $>1,000$ and $<500 \, SD$ (Figure 2a,b). SD $= \text{standard deviation} \times 100$ of the mean monthly temperature). Finally, splits in community composition occurred when sand content was between 45–60%, corresponding to shifts in plant pathogens and saprotrophs (unspecified, dung and soil; Figure 2a,b).

The results of the previous analyses do not depict the relationship between environmental predictors and richness of fungal guilds (they only inform about the existence of a high magnitude change affecting the composition of the community). We, thus, used RF models for each guild and SHAP dependence plots (see Materials and methods) to visualize these relationships (Figure 3).

All these relationships showed different degrees of nonlinear behaviour, with marked shifts in the predictors signalling either abrupt (e.g., changes in DTR, sand, UV or TSEA for unspecified saprotrophs), or nonlinear trends (e.g., changes occurring in TSEA) affecting probability of occurrence of fungal guilds. For instance, plant pathogens and unspecified saprotrophs had a higher probability of occurrence with increases in UV (values $>7.5$), TSEA (values $>900 \, SD$), and decreases in DTR (values $<14 \, ^\circ C$), with pathogens also being positively associated with sand content of approximately 35–50%. Soil saprotrophs were predicted to be more likely to occur with decreasing UV (values $<7.5$), TSEA (values $<500 \, ^\circ C$) and DTR (values $<14 \, ^\circ C$); litter saprotrophs were generally most likely to be more likely to occur at lower TSEA (values $<500 \, SD$).

3.4  |  Global patterns of fungal diversity in drylands

The RF model built to assess the relative contributions of spatial and environmental predictors to overall fungal diversity in drylands revealed a strong contribution of spatial distance (PCNM1) and aridity (sand $= \text{sand content}; \, \text{SOC} = \text{soil organic carbon}$) percentage of increase in mean squared error $>25$ for both: Figure 4a), followed by temperature seasonality and UV index. We observed generally different levels of Shannon diversity among the different aridity classes, with hyperarid, arid and dry sub-humid areas supporting a significantly lower (Wilcoxon test, $p < .05$) fungal diversity than

3.3  |  Detection of community changes

Frequency histograms and density plots of the values used by the classification trees for splits (i.e., the split density plots in Figure 2a) also strongly associating with lichenized fungi ($R^2$ value $= .05$; Supporting Information Figure S8).

FIGURE 1  Environmental predictors of dryland fungal community composition. (a) Relative importance, R-squared ($R^2$), of each environmental predictor included in the gradient forest analysis. (b) Contribution (0 to 1) of climatic, soil and spatial categories to the variation explained by the complete gradient forest model for the 15 ecological groups for which significant predictive power was established ($R$-squared $>0$). PSEA = precipitation seasonality; TSEA = temperature seasonality; DTR = diurnal temperature range; AI = aridity index; UV = UV index; MAT = mean annual temperature; SPEI = standardized precipitation-evapotranspiration index; PCNM1 and 2 = first and second principal coordinates of neighbour matrices; sand = sand content; SOC = soil organic carbon.
semi-arid areas (Figure 4b). Consistently, the fungal maps, estimating the expected geographical distribution and diversity of dryland fungi ($R = .83$, Figure 4c), reflected the extent of well-characterized high classes of aridity. In particular, hotspots of fungal diversity were observed in mid and high latitude drylands (northern Asia, and southern-eastern Europe). Conversely, an overall lower fungal alpha
diversity was projected in mid latitude arid regions (particularly northern Africa and Middle Eastern countries) and the Southern Hemisphere, with the exception of a few hotspots.

4 | DISCUSSION

Our study demonstrates that environmental gradients related to solar UV radiation (i.e., the UV index), climate seasonality (i.e., DTR and TSEA) and soil structure (i.e., sand content) are critical predictors of fungal community changes in global dryland soils, with the greatest influence detected in association with the occurrence of free-living fungi involved in important biogeochemical and biotic processes in dryland soils, such as putative plant pathogens and a range of fungal saprotrophic groups. Further we found that the relationships of environmental predictors with saprotrophic and pathotrophic fungal ecological guilds are markedly nonlinear, exhibiting abrupt shifts in the values of environmental variables that may

![FIGURE 3](https://example.com/figure3.png)

**FIGURE 3** Distribution and environmental predictors of the main fungal ecological groups in drylands. Shapley additive explanations (SHAP) dependence plot of selected climatic and edaphic predictors of plant pathogens and saprotrophs richness in drylands. The effects are expressed as SHAP values, which measure the impact of each predictor on the model output (richness of a particular fungal ecological group). SHAP values are derived for a given predictor value in a process analogous to partial dependence plots; thus, each point on the plot corresponds to a prediction in a sample (see Materials and methods). DTR = diurnal temperature range; UV = UV index; TSEA = temperature seasonality; sand = sand content; R = Spearman’s rho correlation coefficient; p = p value

![FIGURE 2](https://example.com/figure2.png)

**FIGURE 2** Most relevant predictors of fungal composition in drylands worldwide. (a) Frequency histograms of gradient values at which splits occur in the regression trees of the top 15 ecological groups in relationship to the top six environmental variables, showing where along these gradients important compositional changes are taking place. Black lines are the kernel density of the histograms, red lines show the (normalized) distribution of the data along the environmental gradients, and blue lines indicate the ratio between splits and data (ratio between black and red lines). Ratios > 1 (above the dotted line) indicate conditions of relatively greater change in genus composition (i.e., community thresholds). Individual plots depict the predictors, arranged (left to right) from the most to the least important. (b) Compositional change along the top six environmental gradients for the top five fungal guilds for each environmental variable. Each line denotes a guild and their pattern of compositional change along the gradient. The y axes have been normalized so that the maximum corresponds to the relative variable importance. TSEA = temperature seasonality; DTR = diurnal temperature range; UV = UV index; PCNM1 and 2 = first and second principal coordinates of neighbour matrices; sand = sand content
signal particularly vulnerable environmental scenarios. In particular, increases in UV and temperature seasonality above a certain value (7.5 and 900 SD, respectively) were associated with an increased probability of plant pathogen and unspecified saprotroph occurrence, with plant pathogens also being positively associated with sand content of approximately 35–50%. Conversely, these parameters had an overall negative relationship with litter and soil saprotroph incidence, with saprotrophs being negatively influenced also by increases in DTR (values >14°C). Overall, the role of climatic variables observed in our study supports the hypothesis that water-energy dynamics provide an explanation for fungal community turnover across different fungal guilds in drylands across multiple bioregions of the planet.

Mechanistically, the trends we observed at the guild level can be explained by the features that regulate biogeochemical cycling in drylands and the peculiar physiological attributes of saprotrophic and pathogenic fungal guilds (Throop & Archer, 2008). In most arid lands, temperature-related variables and soil structure are considered critical factors in determining the composition of decomposers (Zhu & Cheng, 2011), and traditional models identify extreme temperatures and low soil moisture, typical of dry regions of the world, as major controllers of litter quality and microbial activities (A’Bear et al., 2014). These environmental parameters act as limiting factors for the distribution of fungal decomposers, which tend to withstand overall lower temperature ranges compared to other guilds, such as pathogenic fungi (Větrovský et al., 2019), thus explaining their decrease in occurrence probability with increases in temperature ranges and variability (Feng et al., 2022). Soil structure attributes are also expected to exert various influences on fungal communities, for example by enhancing substrate availability from SOC pools, while also controlling water holding capacity (Hu et al., 2014), which can in turn regulate fungal saprotroph richness and composition. Similarly,
in many arid ecosystems, solar radiation is considered a primary driver of decomposition and carbon cycling (Bornman et al., 2015; Brandt et al., 2010), resulting in a significant photo priming effect that controls root exudation, litter quality and nutrient availability, and accelerates abiotic-driven decomposition in these systems (Day et al., 2018). The tight link between UV radiation and biogeochemical cycling in drylands thus explains the prominent role of the UV index in predicting the distributions of saprotrophic groups associated with soil and litter, and the overall negative influence on soil saprotroph richness.

Conversely, the proportion of putative plant pathogens tended to increase with increasing UV radiation and climate seasonality. This finding is consistent with the physiological traits characteristic of the fungal groups associated with plant-pathotrophic modes (i.e., ascomycetes from the classes Dothideomycetes and Leotiomycetes), which possess resistance to environmental stresses typical of drylands, including UV radiation, high temperature fluctuations, and desiccation (Coleine et al., 2021; Egidi, Delgado-Baquerizo, et al., 2019; Gostincár et al., 2012). Such common adaptations to a wide range of environmental stressors can thus explain the ability of the members of this ecological group to thrive in increasingly extreme environments. Additionally, photoreception and light-dependent traits have been recently suggested as a likely mechanism allowing foliar pathogens from sun-lit habitats to recognize potential partners and stressed hosts (Schumacher & Gorbushina, 2020), indicating that increases in UV radiation might have an important but underestimated role in facilitating the establishment of pathotrophic fungi in dryland ecosystems.

The compositional turnover of the dryland functional mycobiome was also strongly associated with the eigenvector-based spatial descriptors (PCNMs), which were also significantly correlated with the total fungal community diversity. At the guild level, the strongest effect was recorded for pathotrophic and saprotrophic fungi, the most abundant members of the community in our dataset. The large predictive power of PCNMs indicates a role for stochastic processes, such as dispersion limitation, in shaping the community dynamics of the dominant fraction of the fungal assemblies (Dumbrell et al., 2010). Indeed, abundant microbial taxa tend to have higher dispersal rates and to be affected by drift or priority effects more than their rarer counterparts (Mo et al., 2018), possibly explaining the large influence of spatial variables observed in this study for these ecological groups.

The climatic and edaphic parameters considered in our study had poor predictive power for the symbiotrophic guilds in the dataset, such as mycorrhizal, parasitic and endophytic fungi. This might be explained by the typical patchiness of vegetation distribution in drylands (Ding & Eldridge, 2021; Ochoa-Hueso et al., 2018), which might exacerbate the importance of host occurrence for host-associated fungal guilds in dry systems. This finding supports the notion that biotic filters (e.g., vegetation composition; Davison et al., 2021; Hiiesalu et al., 2017), rather than climatic and edaphic factors, might play a critical role in characterizing the distribution of symbiotrophic fungi in dry ecosystems (Tedersoo et al., 2012). For example, ectomycorrhizal fungi are mostly associated with trees in temperate and tropical forests, and might have relatively fewer hosts in arid and hyperarid ecosystems dominated by grasses. However, our study did not include explicit plant predictors, and the relative importance of host and habitat distribution for fungal community turnover in drylands warrants further investigations on a global scale.

Collectively, our results indicate that spatial distance, together with solar UV radiation, temperature and precipitation variability, and soil structure, are underappreciated drivers of global distribution of critically important fungal groups, such as plant pathogens and saprobes, in drylands. These findings imply that processes leading to shifts in UV radiation incidence, soil structure and seasonal climatic patterns, will have disproportionate consequences for the distribution of fungal guilds linked to vegetation and biogeochemical cycling in drylands. These processes might be particularly exacerbated by predicted increases in extreme heatwave events, which could synergistically alter the UV-mediated effects on drylands, thus influencing the balance of plant–soil interactions in these systems (Barnes et al., 2019; Brandt et al., 2007; Vidović et al., 2015). In particular, warming can trigger dramatic water losses in soil that reduce the cooling effect of evaporation and evaporotranspiration, thus enhancing the warming impact of UV solar radiation on the atmosphere (Miralles et al., 2014), with unknown ecosystem-level consequences.

Interestingly, the aridity index, which is considered a primary driver of change in drylands (Berdugo et al., 2020), had a secondary role in determining community turnover in our dataset. However, in line with other global dryland surveys (Maestre et al., 2015), we observed significant decline in fungal alpha diversity with increasing aridity, confirming the critical role of water availability in shaping microbial diversity in global drylands (Maestre et al., 2021). In particular, our predictive maps indicate higher fungal diversity in high latitude areas, which support mostly temperate and cold drylands. On the contrary, lower latitudes, which support mostly arid and hyperarid hot deserts, harboured a comparatively lower projected diversity of fungi. Consistently, temperature seasonality was the second most important environmental explanatory variable associated with the measured diversity of fungi. Similar patterns have been observed also in global surveys of fungi from forest-dominated systems (Větrovský et al., 2019), possibly indicating that aridity and temperature interact to shape fungal diversity on a global scale. Consequently, the diversity of dryland fungi might be particularly vulnerable to the effect of global warming and the magnitude of this effect could be especially important at more extreme latitudes, as recent studies predict accelerated increases in temperatures and temperature extremes towards the poles (Shi et al., 2021).

Taken together, the comprehensive catalogue of ecology–climate relationships we provide paves the way to a more exhaustive and detailed understanding of the complex role of climate and soil in regulating fungal biogeography, especially in those regions of the world that are most vulnerable to environmental changes, such as global drylands. It is important to note, however, that the fungal surveys included in our study had an overall patchy distribution and were biased towards arid to dry sub-humid drylands from temperate...
and cold regions of the American, European and Asian continents, while having poor coverage of potentially important regions, such as the Sahara, Arabian and Australian sandy deserts. It will be thus important in the future to validate our projections on fungal diversity in these undersampled regions of the world. Similarly, the lack of soil data from Antarctica in global databases did not allow us to make predictions of fungal diversity and community turnover in that continent. Further biodiversity inventories including regions of the world systematically excluded from microbial and environmental surveys are required to obtain a full-scale assessment of the diversity and composition of soil fungi in drylands on a global level. It is also important to note that the climatic grids (~1 km) used here and in most global microbial biogeography studies can only allow for coarse-scale evaluations of the association between environmental variables and microbial communities, especially considering that the dependence of the climatic controls on microbial community composition might decrease with depth (Dove et al., 2021). While large scale studies are critical to improve our understanding of the global distribution of multiple soil communities and their possible future trends, efforts are necessary to reduce these uncertainties and increase confidence in mapping microbial communities and predicting where change may be important at national to regional scales.

Finally, we observed a large proportion of fungal taxa unclassified at the guild level, indicating that a substantial effort is still required to obtain a comprehensive overview of the fungal ecological communities inhabiting drylands worldwide, and validate their relationship with environmental change. Our work opens a new line of investigation to include quantifying the importance of ecological processes that govern fungal communities across contrasting regions of the world, with particular emphasis on identifying the traits, and trait trade-offs, underpinning their dispersal and functional capabilities in such unique ecosystems (Zanne et al., 2020). We anticipate that, as strain-specific trait data become available and ecological guild assignments more comprehensive, better assessment of functional variation expressed within and among communities in relation to dispersal constraints, UV tolerance and climate variability could be performed. This information is required to provide better predictions of the current and future adaptation of fungi to the effects of climate change, and possible ramifications for the sustainability of dryland ecosystems.

ACKNOWLEDGMENTS
C.C. is supported by the European Commission under the Marie Skłodowska-Curie Grant Agreement No. 702057 (DRYLIFE). C.C. acknowledges funding from the Italian National Program for Antarctic Research (PNRA) and is supported by a PNRA postdoctoral fellowship. E.E. is supported by an Australian Research Council DECRA (Discovery Early Career Researcher Award) fellowship (DE210101822). M.D-B. is supported by a project from the Spanish Ministry of Science and Innovation (PID2020-115813RA-I00), and a project PAIDI (Andalusian Research, Development and Innovation Plan) 2020 from the Junta de Andalucía (P20_00879). Microbial distribution and colonization research in B.K.S.’s lab is funded by the Australian Research Council (DP210102081). E.G. is supported by the European Research Council Grant agreement 647038 (BIODESERT).

CONFLICT OF INTEREST
The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT
Raw data details are reported in Supporting Information Appendix S1 – Data Sources. Metadata are available in Supporting Information Table S1. The codes for the computational analyses are available in Figshare (https://doi.org/10.6084/m9.figshare.19243749.v1).

CLINICAL TRIAL REGISTRATION
Not applicable.

ETHICS APPROVAL STATEMENT
Not applicable.

PATIENT CONSENT STATEMENT
Not applicable.

PERMISSION TO REPRODUCE MATERIAL FROM OTHER SOURCES
Not applicable.

ORCID
Eleonora Egid
https://orcid.org/0000-0002-1211-2355
Manuel Delgado-Baquerizo
https://orcid.org/0000-0002-6499-576X
Miguel Berdugo
https://orcid.org/0000-0003-1053-8907
Emilio Guirado
https://orcid.org/0000-0001-5348-7391
Davide Albanese
https://orcid.org/0000-0002-9493-3850
Brajesh K. Singh
https://orcid.org/0000-0003-4413-4185
Claudia Coleine
https://orcid.org/0000-0002-9289-6179

REFERENCES


Supporting Information

Additional supporting information can be found online in the Supporting Information section at the end of this article.