Modeling geographic distribution of arbuscular mycorrhizal fungi from molecular evidence in soils of Argentinean Puna using a maximum entropy approach

Davide Nepote Valentin [Equal first author, 1] , **Samuele Voyron** [Equal first author, 1, 2] , **Florencia Soteras** [3] , **Hebe Jorgelina Iriarte** [4, 5] , **Andrea Giovannini** [1] , **Erica Lumini** [Corresp., 2] , **Mónica A. Lugo** [4, 5]

[1] Life Sciences and Systems Biology, University of Turin, Torino, Italy

[2] Institute for Sustainable Plant Protection (IPSP), National Research Council (CNR), Torino, Italy

[3] Laboratorio de Ecología Evolutiva y Biología Floral, Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET, FCEFyN, Universidad Nacional de Córdoba, Córdoba, Argentina

[4] Instituto Multidisciplinario de Investigaciones Biológicas (IMIBIO-CONICET-UNSL), San Luis, Argentina

[5] Micología, Diversidad e Interacciones Fúngicas (MICODIF), Área Ecología, Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis (UNSL), San Luis, Argentina

Corresponding Author: Erica Lumini
Email address: erica.lumini@ipsp.cnr.it

The biogeographic region of Argentinean Puna mainly extends at elevations higher than 3000 m a.s.l. within the Andean Plateau and hosts diverse ecological communities highly adapted to extreme aridity and low temperatures. Soils of Puna are typically poorly evolved and geomorphology is shaped by drainage networks, resulting in highly vegetated endorheic basins and hypersaline Salar basins. Although plant ecology has been investigated for decades, Fungi distribution in the soils of Puna still lacks detailed research. We collected soil samples from 28 points, following a nested design within three different ecological environments, i.e. grazed or undisturbed shrublands and grasslands, a hypersaline Salar basin area, and family-run crop fields. Total Fungi and Arbuscular Mycorrhizal Fungi (AMF) occurrence was investigated through eDNA sequencing. Furthermore, the importance of soil chemical parameters and environmental variables in shaping AMF presence was assessed through a Generalized Mixed Model and Multi Model Inference approach, highlighting the relevance of soil salinity and organic matter content as significant predictors of AMF occurrence in dry Puna soils. It was also estimated if intensive grazing by cattle and lama may reduce the presence of AMF in these highly disturbed soils, driving major ecological changes in punean grasslands. Species Distribution Modelling was performed within an environmental coherent area comprising both phytogeographic regions of Puna and Altoandino above 27° latitude S. We modeled AMF distribution with a maximum entropy approach, using both bioclimatic and edaphic predictors, and plotted maps of environmental idoneity of the predicted area for AMF. These approaches confirmed that lower suitability for AMF is predicted in hypersaline Salar

areas, while grassland habitats and a wider temperature seasonality range appear to be related to AMF enrichment, suggesting a main role of seasonal dynamics. The highest abundance of AMF was however observed in Vicia faba crop fields while potato fields yielded a very low AMF occurrence. Given the role of arbuscular mycorrhization in enhancing plant growth dynamics and soil fertility, we set a new series of models excluding the cultivated Chaupi Rodeo samples. These models were run to assess the impact of farming, and highlighted that if these areas remained unmanaged habitats of Puna and Altoandino then, large-scale soil features and local bioclimatic constraints would likely support a lower idoneity for AMF.

1 **Modeling geographic distribution of arbuscular mycorrhizal**
2 **fungi from molecular evidence in soils of Argentinean Puna**
3 **using a maximum entropy approach**
4
5 Davide Nepote Valentin [1,§], Samuele Voyron [1,2§], Florencia Soteras [3], Hebe J. Iriarte [4,5], Andrea
6 Giovannini [1], Erica Lumini [2*],Mónica A. Lugo[4,5]
7
8 [1] Department of Life Sciences and Systems Biology, University of Torino, Torino, Italy.
9 [2] Institute for Sustainable Plant Protection (IPSP), National Research Council (CNR), Torino,
10 Italy
11 [3] Laboratorio de Ecología Evolutiva y Biología Floral, Instituto Multidisciplinario de Biología
12 Vegetal (IMBIV), CONICET, FCEFyN, Universidad Nacional de Córdoba, Argentina
13 [4] Instituto Multidisciplinario de Investigaciones Biológicas (IMIBIO-CONICET-UNSL), San
14 Luis, Argentina
15 [5] Micología, Diversidad e Interacciones Fúngicas (MICODIF), Área Ecología, Facultad de
16 Química, Bioquímica y Farmacia, Universidad Nacional de San Luis (UNSL). San Luis,
17 Argentina
18
19 § these authors contributed equally
20
21 *Corresponding Author:
22 Erica Lumini
23 Viale PA Mattioli 25, 10125 Turin, Italy,
24 Email address: erica.lumini@ipsp.cnr.it
25

26 **Abstract**
27 The biogeographic region of Argentinean Puna mainly extends at elevations higher than 3000 m
28 a.s.l. within the Andean Plateau and hosts diverse ecological communities highly adapted to
29 extreme aridity and low temperatures. Soils of Puna are typically poorly evolved and
30 geomorphology is shaped by drainage networks, resulting in highly vegetated endorheic basins
31 and hypersaline *Salar* basins. Although plant ecology has been investigated for decades, Fungi
32 distribution in the soils of Puna still lacks detailed research.
33 We collected soil samples from 28 points, following a nested design within three different
34 ecological environments, i.e. grazed or undisturbed shrublands and grasslands, a hypersaline
35 Salar basin area, and family-run crop fields. Total Fungi and Arbuscular Mycorrhizal Fungi
36 (AMF) occurrence was investigated through eDNA sequencing. Furthermore, the importance of
37 soil chemical parameters and environmental variables in shaping AMF presence was assessed
38 through a Generalized Mixed Model and Multi Model Inference approach, highlighting the
39 relevance of soil salinity and organic matter content as significant predictors of AMF occurrence
40 in dry Puna soils. It was also estimated if intensive grazing by cattle and lama may reduce the
41 presence of AMF in these highly disturbed soils, driving major ecological changes in punean
42 grasslands. Species Distribution Modelling was performed within an environmental coherent
43 area comprising both phytogeographic regions of *Puna* and *Altoandino* above 27° latitude S. We
44 modeled AMF distribution with a maximum entropy approach, using both bioclimatic and
45 edaphic predictors, and plotted maps of environmental idoneity of the predicted area for AMF.
46 These approaches confirmed that lower suitability for AMF is predicted in hypersaline Salar
47 areas, while grassland habitats and a wider temperature seasonality range appear to be related to
48 AMF enrichment, suggesting a main role of seasonal dynamics. The highest abundance of AMF
49 was however observed in *Vicia faba* crop fields while potato fields yielded a very low AMF
50 occurrence. Given the role of arbuscular mycorrhization in enhancing plant growth dynamics
51 and soil fertility, we set a new series of models excluding the cultivated *Chaupi Rodeo* samples.
52 These models were run to assess the impact of farming, and highlighted that if these areas
53 remained unmanaged habitats of Puna and Altoandino then, large-scale soil features and local
54 bioclimatic constraints would likely support a lower idoneity for AMF.
55

56 **INTRODUCTION**
57 The South American biogeographic region of Puna extends within the Andes Mountain Range
58 across Argentina, Bolivia, Perú and Chile at elevations higher than 3000 m a.s.l. In the Quechua
59 local native language Puna means "high and cold lands" (Carilla et al., 2018), which meaning
60 well matches the wide extent of biogeographical, ecological and geographical Puna concepts.
61 These high Andean plateau regions encompass diverse ecological communities which share
62 severe or extreme aridity and wide seasonal and daily temperature variations (Lugo & Menoyo,
63 2019 and references therein). In the Argentinean Puna, the Puna geological province (Turner &
64 Méndez 1979) is the geomorphological basement on which the Puna region (Morrone, 2001), the
65 Puna and Altos Andes ecoregions (Matteucci, 2018a; Matteucci, 2018b), the phytogeographic
66 provinces of the Dry, Humid and Desert Puna (Troll, 1959; Troll, 1968), the phytogeographic
67 provinces Puneña and Altoandina (Cabrera, 1976; Cabrera & Willink, 1980; Carilla et al., 2018;
68 Oyarzabal et al., 2018) and Puna floristic districts "Jujeño", "Central", "Cuyano", and
69 "Boliviano" (Martínez Carretero, 1995) are placed. All these floristic, biogeographical and
70 ecological ecosystems of the Puna biome exceed the limits of the Puna geological province
71 located at Catamarca Province (Turner & Méndez, 1979) and extend south to the Mendoza

72   Province (Martínez Carretero, 1995, Matteucci, 2018b) in Argentina. Immersed in this common
73   geographic foundation of the Puna geological province with its different plant communities,
74   there is a remaining strip of the Altoandina phytogeographic province included in the Puneña
75   phytogeographic province (Cabrera 1976; Cabrera & Willink, 1980; Carilla et al., 2018;
76   Oyarzabal et al., 2018), and a similar areal intrusion also occurred for Altos Andes and Puna
77   ecoregions (Matteucci, 2018a; Matteucci, 2018b); due to the fact that these phytogeographic
78   provinces and ecoregions are extending along similar territories, hereafter these are named Puna
79   and Altoandino, respectively. The above-mentioned strip area of Altoandino is represented by
80   Chaupi Rodeo (Jujuy, Argentina), one of the sampled sites of this work where a native people's
81   settlement and their crops are placed. Puna and Altoandino are discriminated mainly by their
82   distinct elevations, reaching 3,000-3400 m, and 4,000- 4,500 m, respectively (Carilla et al., 2018;
83   Cabrera, 1976; Oyarzabal et al., 2018). Despite the different altitudinal range and its effect on
84   vegetation features, physiognomic, phytosociological, and differing ecological categorizations
85   between Puna and Altoandino regions of Northwest Argentina, similar geological, historical and
86   ecological processes suffered along the time (Matteucci, 2018a; Matteucci, 2018b) led some
87   authors to consider them as a unique and homogeneous area. These specific environmental
88   conditions make this area an ideal region for studying the distribution patterns of AMF in
89   relation to abiotic variables. Puna and Altoandino host shrub-dominated vegetation units and
90   gramineous steppe, as well as other azonal communities such as halophyte within salt flats, the
91   hypersaline endorheic basin named Salina or Salar areas, and Poaceae, Juncaceae, Cyperaceae
92   within Vegas wet flooded endorheic basins; however, vegetation cover is not homogeneous
93   within the same phytogeographical province or ecoregion (Martínez Carretero, 1995; Renison et
94   al., 2013; Carilla et al., 2018; Matteucci, 2018a; Matteucci, 2018b; Oyarzabal et al., 2018).
95   Soils of Puna and Altoandino are typically poorly evolved, influenced by cryogenic processes
96   and aridity (Panigatti, 2010). These ecosystems are highly exposed to desiccation, extreme
97   environmental conditions such as large daily temperature amplitudes, an incident solar energy
98   greater than 2200 KW/m2/year and intense UV irradiation (Martínez Carretero 1995; Carilla et
99   al., 2018; Matteucci, 2018a; Matteucci, 2018b). These harsh conditions result in high community
100 vulnerability when exposed to sudden ecological changes, mainly caused by the anthropic impact
101 of extensive grazing of camelids and cattle, wildfires, mining, wood harvesting (Carilla et al.,
102 2018; Matteucci, 2018a; Matteucci, 2018b), and consequently, increasing desertification
103 processes at an extremely high level for these ecosystems (Vorano & Vargas Gil, 2002). With
104 regard to extensive grazing, livestock management in these areas usually follows the nomadic
105 type with continuous migration of multispecies herds made up of sheep, goats and llamas
106 (Vorano & Vargas Gil, 2002; Quiroga Mendiola & Cladera, 2018). Particularly, flocks of native
107 camelids are represented by two wild species (huanaco, *Lama guanicoe* Statius Müller and
108 vicuña, *Vicugna vicugna* Molina) and two domesticated species (llamas, *Lama glama* L. and
109 alpacas, *Vicugna pacos* L.) but only in few areas there is cows and camelid flocks circulation,
110 and in some cases llamas or llamas and cattle, are restricted by paddocks or corrals (Carilla et al.,
111 2018). This migratory grazing system likely turned natural grasslands into shrublands.
112 Furthermore, this closure-based grazing system proved to have a greater impact on the vegetation
113 than the extensive grazing by native camelids (Quiroga Mendiola & Cladera, 2018).
114 Environmental exploitation impacts on soil fertility and ecosystem services that local
115 subsistence-based native communities rely on, with direct effects on their economy and living
116 conditions. In this perspective, investigating the microbiological diversity of these soils might be
117 a valuable indicator of the ecological conditions and of the disturbance of these ecosystems.

118   Soil-borne fungi play a fundamental role in the ecosystem functioning, as decomposers of
119   organic matter, inducing soil aggregation, as pathogens, or as mutualists assisting plants growth.
120   Among soil fungal communities, arbuscular mycorrhizal fungi (AMF) are worldwide obligate
121   symbionts belonging to Phylum Glomeromycota (Wijayawardene et al., 2020), that associate
122   with plants and promote their growth by improving soil nutrients and water uptake as well as
123   providing pathogen protection to their host (Smith & Read, 2008). Despite the AMF widely
124   assessed ubiquity, their occurrence may be affected by environmental conditions (e.g. soil
125   nutrients, pH, precipitation, temperature) and vegetation type (Veresoglou et al., 2013; Davison
126   et al., 2021). For instance, in a global analysis Tedersoo et al., (2014) evidenced a positive
127   relationship of AMF richness with potential evapotranspiration and soil pH, as well as a diversity
128   increase in grassland and shrubland ecosystems. Therefore, AMF richness and distribution is
129   expected to be affected by current land uses, such as cropping and grazing, and the study of the
130   AMF community patterns in relation to climatic and edaphic conditions may help to predict their
131   response to global change (Kivlin et al., 2017). Species distribution modeling (SDM) combines
132   occurrence data with environmental variables to geographically predict potential suitable areas
133   within the studied environments (Elith et al., 2009). This approach has been rarely used for
134   estimating AMF distribution (Kivlin et al., 2017), mainly because of the difficulty in delimiting
135   the potential area of occurrence for these soil-borne organisms. However, considering that
136   geological, environmental and soil conditions have delimited the geographical boundaries of our
137   study area, the ecosystems of Puna include promising regions to predict AMF distribution.
138   Mycorrhizal fungi are a main driver of soil nutrients and plant-soil ecology (Wurzburger et al.,
139   2017), thus representing an ideal target for a microbiological comparative analysis of soils from
140   different habitats within Puna. Considering the paucity of research addressing the fungal profile
141   of Puna soils (Lugo et al., 2008, Ontivero et al., 2020), and the importance of AMF in plant-soil
142   dynamics (Davison et al., 2021, Dumbrell et al., 2010; Bonfante & Genre, 2015), in the present
143   study we focused on this fungal group, relying on the fungal OTU table provided in Ontivero et
144   al., (2020) and extending the analysis to new soil sampling points from grazed and undisturbed
145   areas of Puna and from an endorheic Salar basin area within Puna. We first estimated the relation
146   of environmental predictors with AMF occurrence by model weighting and averaging. Then, we
147   performed a maximum entropy approach for SDM (Bradie & Leung, 2017; Zimmermann et al.,
148   2010; Austin, 2002) in order to predict the environmental suitability for AMF in the Argentinean
149   Puna and Altoandino, phytogeographically and ecoregionally delimited as mapped by Oyarzabal
150   et al., (2018), bounding it above 27° latitude S. To maintain a relatively homogeneous area of
151   prediction and being both above indicated areas originated by a unique geomorphological event
152   (Matteucci, 2018a; Matteucci, 2018b), we considered them as a unique geological province
153   named Puna (Turner & Méndez, 1979). Since Chaupi Rodeo is located in a fringe of Altoandino
154   extending into Puna and the current human agricultural use of soil in this area might have
155   influenced AMF communities, we tested different models, including or not the Chaupi Rodeo
156   sampling points, and compared the differential suitability of the predicted area for AMF.

157   ## MATERIALS & METHODS

158   **Soil sampling and environmental metadata**

159   The sampling sites can be mostly included in Puna, and more specifically in the so-called Dry
160   Puna (Troll, 1959; Troll, 1968) because their annual rainfall ranges between 100 and 400mm,
161   and are restricted to the summer season with consequent very dry and cold winters. Although

162 Chaupi Rodeo sampling site presents the same climatic conditions of the other sampled areas, its
163 geographical placement overlaps Altoandino boundaries. Soil samples have been collected from
164 28 sampling points along transects within six different locations in northern Argentinean Jujuy
165 Province belonging to three main different environments: (i) habitats of Puna grazed by lama or
166 lama and cattle, or undisturbed for the locations named Dunas, Punto Susques, Puesto del
167 Marqués, Abra Pampa, (ii) hypersaline Salar basin area for the location Salinas Grandes, and (iii)
168 family-run crop fields for three Chaupi Rodeo locations in Altoandino (Ontivero et al., 2020),
169 from now on named Chaupi Rodeo A, B and C (Fig.1). Multiple environmental predictors have
170 been considered per each sampling point, namely elevation, grazing type, preceding habitat or
171 crops in the previous year and plant cover. Soil physicochemical analyses were also carried out
172 in the Soil Laboratory of INTA, EEA Villa Mercedes, San Luis, Argentina. Each soil sample was
173 analyzed for pH, percentage of organic matter (Walkley Black method), soil carbon quantity
174 (Walkley Black method), available phosphorus (Bray and Kurtz method) and electric
175 conductivity as a proxy for salinity. Precise geolocation for each sampling point was recorded.

176 **Sequencing and bioinformatic analysis**

177 DNA extraction, PCR amplification, amplicons sequencing and bioinformatic analysis have been
178 performed as described in Ontivero et al., (2020). Soil samples were sieved through a 2-mm
179 mesh size sieve. Genomic DNA was extracted from 250 mg of each soil sample by means of the
180 DNeasy PowerSoil kit (Qiagen, CD Genomics Company, Shirley, NY 11967, USA) according to
181 the manufacturer's instructions. In order to investigate, not only the AMF biodiversity but also
182 the total fungal communities, the ITS2 region, together with the AMF-specific barcode region,
183 were amplified by a nested approach based on the protocol described by Berruti et al., (2017).
184 The obtained PCR products were checked on 1% agarose gel, purified by means of Wizard SV
185 Gel and PCR CleanUp System (Promega, Wisconsin, Madison, USA) quantified with Qubit 2.0
186 (Thermo Fisher Scientific, Waltham, MA, USA) and sent to BMR Genomics (Padova, Italy) for
187 Illumina MiSeq platform (2 × 300 bp) sequencing. For each sampling point, reads were
188 sequences with forward and reverse reads in separate files. For all datasets, sequencing adapters
189 and primers were removed, and the sequences were then analyzed by means of the microbiome
190 bioinformatics platform QIIME2 (Quantitative Insights Into Microbial Ecology 2) version
191 2019.7 (Bolyen et al., 2019). Denoising and quality control, including removal of chimeras, were
192 achieved by means of the DADA2 (Callahan et al., 2016) plugin (qiime dada2 denoise-paired)
193 and, to avoid low-quality sequences, reads were truncated (>280 bp for forward, >265 bp for
194 reverse reads). Feature tables for AMF were generated by means of qiime vsearch cluster-
195 features-de-novo plugin using 97% as the identity threshold. The classifier adopted for the
196 taxonomic assignment of the total fungal community was generated using the UNITE
197 Community (2019): UNITE QIIME release for Fungi version 10.05.2021 (Abarenkov et al.,
198 2021). Ecological roles of the fungal communities at each sampled site were inferred using
199 FUNGuild (Nguyen et al., 2016). Finally, the taxonomy of AMF OTUs was refined, performing
200 BLAST against the MaarjAM database (https: //maarjam.botany.ut.ee/; pik et al., 2010). The data
201 sets, composed of OTU table, taxonomy table and metadata were then imported in Rstudio and
202 were used to create two phyloseq objects with the R package phyloseq (McMurdie and Holmes,
203 2013) that were employed for all the following analyses. Krona plots were generated using
204 Krona Tools and cpauvert/psadd library (Ondov et al., 2011). The dataset generated for this
205 study can be found in the NCBI Sequence Read Archive (SRA-NCBI; https://www.ncbi.
206 nlm.nih.gov/sra) under project accession number PRJNA835719.

GLMM and model averaging

To analyze the AMF communities, the OTU (Operational Taxonomic Units) table was rarefied at an even sequencing depth of 19,882 sequences per sample and restricted to the Phylum Glomeromycota. Due to a non-normal distribution of data, a non-parametric Kruskal-Wallis analysis of variance was performed to test for the difference among locations using the function kruskal.test in the R stats package version 4.1.0 (R core Team, 2013). Post-hoc pairwise tests were then performed applying a Bonferroni correction by using the function pairwise.t.test in the R Stats package. Sample means were then grouped by location and plotted by each environmental predictor as listed above, using the R package ggplot2 version 3.3.5 (Wickham, 2016). Because of the nestedness of the sampling design and the overdispersion of data, a negative binomial GLMM (Generalized Linear Mixed Model) approach was then performed, using the function glmer.nb in the R package blmeco version 1.4. (Korner-Nievergelt et al., 2015), in order to account for any possible random effect of samples and locations. Collinearity among predictors was previously tested by means of Variance Inflation Factor approach using the R package car version 3.0-12 (Fox & Weisberg, 2019). A Pearson's r coefficient value of 0.7 was chosen as a collinearity threshold. The model selection was performed both through a Minimum Adequate Model and a Multi Model Inference approach using the R function drop1 and the R function dredge in the package MuMIn version 1.43.17 (Barton, 2020). The sum of weights for each predictor was estimated from the best competing models, then model-averaged coefficients were estimated and the difference from zero tested for significance. Beta-diversity among different locations and among habitats was assessed by means of Non-metric Multi-Dimensional Scaling (NMDS) ordination system using Jaccard dissimilarity index and statistical significance was verified performing a PERMANOVA test, using the software MycrobiomeAnalist (Chong et al., 2020; Dhariwal et al., 2017).

Geographical Information Systems and Species Distribution Models

A Species Distribution Model approach was used to model the potential distribution of AMF by estimating the environmental suitability through the maximum entropy-based software MaxEnt (Phillips et al., 2004 2006, Elith et al., 2011, Warren & Seifert, 2011) version 3.4.1. A relatively biogeographically homogeneous area was selected to be modeled, following the boundaries of phytogeographic units of Puna and Altoandino as mapped in Oyarzabal et al., (2018). Despite the site Chaupi-Rodeo is located below 4000 m.a.s.l and therefore it would not be strictly within the Alto-andina unit definition of Oyarzabal (2018), this site falls within the boundaries of this phytogeographic unit. Therefore, and considering that Chaupi-Rodeo sampling points also fall within the Altoandino or Altos Andes ecoregion sensu Matteucci, 2018a, we included Alto-andina region within the predicted area. Due to similar tectonic and orogenic dynamics, Argentinean Puna and Altoandino can be considered as a unique region northern of 27° latitude S, therefore allowing to model the environmental suitability for both. Consequently, the estimated area was bounded southward by this latitude (Matteucci, 2018a; Matteucci, 2018b) to avoid excessive environmental heterogeneity for the predictions. Cartographic processing was performed through the software QGIS version 3.16 (QGIS Development Team, 2021) using the World Geodetic System WGS84 for rasters and vectors georeferencing. A set of climatic and environmental rasters were used as predictors of the suitability of AMF. Layers were previously clipped within the selected boundaries, and aligned to fit the georeferencing. Nineteen Worldclim 2 (Fick & Hijmans, 2017) bioclimatic variables regarding temperature and precipitation and elevation at maximum resolution of 30" were chosen as predictors along with

252 the following environmental rasters: Land Cover at 30" resolution by National Mapping
253 Organizations - GLCNMO based on index LCCS developed by FAO; vegetation units
254 (Oyarzabal et al., 2018) corresponding to Puna (dominated by *Fabiana densa* Remy and
255 *Baccharis boliviensis* (Wedd.) Cabrera and Altoandino (dominated by *Senecio algens* Wedd. and
256 *Oxalys compacta* Gillies ex Hook. & Arn.); soil nutrients retention and soil nutrients availability
257 by FAO Geonetwork, and soil type by SMW-Digital Soil Map of the World - FAO (Fig. 2). The
258 selection of climatic variables was performed stepwise and climatic variables collinearity was
259 tested by means of the R package corrplot version 0.92 (Wei & Simko, 2021) and stepwise
260 Multivariate Principal Components Analysis (PCA) were performed on the bioclimatic values
261 matrix using the software Past 4 (Hammer et al., 2001). A first set of models were generated
262 through MaxEnt using the 19 Worldclim bioclimatic variables. Permutation importance values
263 and jackknife AUC (area under the receiver operating curve) graphs were cross-referenced with
264 PCA plots to perform a final bioclimatic variables selection. The selected bioclimatic variables
265 were added to the above described environmental and edaphic predictors and a new set of 10
266 bootstrap MaxEnt models was generated, setting the test data to 30% and optimizing the process
267 to maximum 5000 iterations and 10,000 background points of pseudo-absence. A last stepwise
268 process of variable selection was then performed running multiple models in order to select the
269 best model among those generated, by comparing AUC values, permutation importance of
270 variables and jackknife AUC test results. A presence-absence threshold was then applied both to
271 cloglog and cumulative outputs of the model, generating QGIS maps to show the predicted areas
272 of presence and of absence. The presence-absence cumulative threshold was calculated as a
273 mean of fixed cumulative 10% values of threshold generated by MaxEnt for each of the 10
274 bootstrap model replicates, while the cloglog threshold was calculated as a mean of the
275 thresholds generated by MaxEnt by balancing training omission, predicted areas and threshold
276 values. Omission rates were calculated per each threshold and tested for significance. Since all
277 the Chaupi Rodeo points of presence were sampled in agricultural soils, while the other samples
278 were collected in unfarmed environments of Puna and Salar, the above-described modeling
279 procedure was performed again on a subset of points of presence excluding the Chaupi Rodeo
280 location by using the same variables selected for the full points model. This procedure not only
281 accounts for the possible biases in predicting AMF suitability when considering both cultivated
282 and uncultivated soils but also allows a direct comparison between the potential distribution
283 predicted using all points and an unmanaged environment-based one.

284 **RESULTS**

285 **Total fungal abundances and beta-diversity**

286 The taxonomic analysis of the retrieved fungal communities showed that the highest number of
287 reads belonged to the Phylum Ascomycota with more than 30,000 sequences per location,
288 followed by Phylum Basidiomycota, ranging from more than 8,500 up to almost 32,000
289 sequences across samples (Fig. 3, TableS1). The Phylum Glomeromycota showed a high
290 abundance per sample above 900 sequences only in the croplands of Chaupi Rodeo, with the
291 exclusion of potato fields, and in the sampling point Cal_D in Abra Pampa (Fig. 3, TableS2).
292 Read counts higher than 300 sequences were also observed for the Phylum Mortierellomycota in
293 Chaupi Rodeo B and C locations and in Salinas Grandes soils (Fig. 3), while all other Phyla
294 showed read abundances lower than 100 sequences. Beta-diversity among different locations was
295 assessed by means of Non-metric Multi-Dimensional Scaling (NMDS) ordination system using

296   Jaccard dissimilarity index and statistically tested by means of PERMANOVA, with a R-squared
297   value of 0.41928 and p-value < 0.001 confirming significant differences among locations,
298   Salinas Grandes samples clearly separated from the other locations (Fig. 4). Significant
299   differences in Beta-diversity among habitats were also assessed, with a R-squared value of
300   0.26045 and a p-value < 0.001, thus confirming differences among the considered sampling
301   points in Puna, Salar and Chaupi Rodeo different crops.
302

Occurrence of Phylum Glomeromycota and ecological drivers

304   Kruskal-Wallis non-parametric test of variance showed an overall significant difference in AMF
305   abundance among locations (chi-squared=54.393, p-value=1.966e-09; Fig. S1) and single
306   samples (chi-squared=106.21, p-value=2.411e-11). . As a matter of fact, a very scarce
307   occurrence of AMF was observed in the surrounding areas of Salar (Dunas) where we found
308   Glomeromycota sequences only in one sample, while no sequences were found in Salinas
309   Grandes soil samples nor in two of the three potato fields in Chaupi Rodeo (CRP_B and
310   CRP_C). Significant differences were also observed among habitats, i.e. Puna, Salar, corn crops,
311   *Vicia faba* L. crops and potato crops (chi-squared 37.742, p-value 1.266e-07; Fig. 5), as well as
312   among types of soil anthropic impact, i.e. lama or lama and cattle grazed areas, undisturbed or
313   farmed soils (chi-squared =27.814, p-value =3.973e-06; Fig. S2.A) and among samples with
314   different land uses in the previous year (chi-squared= 20.993, p-value = 0.0003177; Fig. S2.B).
315   Pairwise post-hoc tests with Bonferroni correction (Table S3) revealed that Chaupi Rodeo
316   cropland soils were significantly different from the soils of the areas of Salinas Grandes and
317   Dunas, respectively extracted from a Salar and from less saline surroundings of the Salar. Only
318   one of the three locations in Chaupi Rodeo (Chaupi Rodeo_A) was significantly different from
319   the area of Punto Susques, located within the Puna habitat. The above stated significances among
320   locations were due to the samples CRF_A and CRF_B, both *V. faba* crops in Chaupi Rodeo,
321   which were significantly different from all Dunas and Salinas Grandes samples, two Chaupi
322   Rodeo potato crops (CRP_B and CRP_C) and samples PdM B and PS_C, these last two both
323   belonging to Puna habitats. Significant differences in Glomeromycota reads were also observed
324   among cultivated soils and soils of Puna grazed or not by Lama and among soils of crops grown
325   at *V. faba* and all the other soils as well as among corn crops and Puna and Salar soils.
326   Adding soil biochemical parameters as predictors of AMF occurrence in a Generalized Linear
327   Mixed Model approach assessed soil salinity and the quantity of organic matter in soil as
328   significant drivers of AMF occurrence in the sampled areas (Table S4). Salinity showed the
329   biggest size effect and a negative correlation thus reflecting how increasing salinity in the
330   investigated soils seems to limit AMF occurrence while increasing organic matter may benefit
331   AMF growth. The nested random effects of samples within locations showed variance values
332   greater than zero, hence accounting for the spatial autocorrelation of data. Consistent results
333   were obtained by a Multi Model Inference (MMI) approach. The sum of weights of predictors
334   calculated from model weighting and the model-conditional averaged coefficients highlighted
335   how salinity and organic matter appear to be the more influencing predictors on AMF occurrence
336   in the investigated soils, with estimates significantly different from zero (Table 1).
337

Modeling AMF geographical and environmental suitability in Argentinian Puna
339   To model the distribution of AMF in the selected areas of Puna and Altoandino, a stringent
340   stepwise climatic and environmental variables selection through an iterative maximum entropy

341   modeling approach, was carried out as described above. Principal Component Analysis (PCA)
342   performed on bioclimatic and elevation values, extracted from each point of AMF presence,
343   allowed to select separately a subset of temperature and precipitation variables likely explaining
344   more variance to complement the importance percentages assigned by the software MaxEnt to
345   variables (Fig. S3). The selected bioclimatic variables chosen to run the final set of models were
346   minimum temperature of coldest month (bio6), temperature annual range (bio7), and coefficient
347   variation of precipitation seasonality (bio15), which were added to the environmental predictors
348   from the selected rasters as described above. Among a set of competing models with AUC values
349   larger than 0.975 generated using MaxEnt, a best predictive model was chosen scoring a value of
350   AUC of 0.978± 0.005 for the cloglog output and a value of AUC of 0.978 ± 0.007 for the
351   cumulative output. MaxEnt permutation importance of variables values (Table S5) revealed that
352   the variables with a larger contribution in explaining the model were temperature annual range
353   (bio7), soil type and land cover, with lower contributions given by elevation, nutrient retention in
354   soils, quantity of nutrients, vegetation type, and precipitation seasonality (bio15). An overall
355   assessment of the cloglog and cumulative models can be inferred observing the plots showing the
356   training omission rate and the predicted area as a function of the cumulative threshold (Fig. 6),
357   averaged over the replicate runs, here complementing the AUC value and confirming a viable
358   quality of the models for the predicted area, notably considering the limited number of presence
359   points derived from the sampling design. Cloglog and cumulative suitability maps of AMF
360   occurrence were generated as an output of the MaxEnt runs, respectively following a probability
361   and a percentage scale. As expected, cumulative prediction expands the suitability to larger areas
362   than the cloglog model (Fig. 7).
363   A threshold value was chosen among those calculated by MaxEnt for each of the ten model
364   repetitions. Only thresholds corresponding to low or null omission rates and tested for
365   significance (p-value<0.05) were considered as possible presence-absence cutoffs. The MaxEnt
366   threshold generated as a balance among training omission, predicted area and threshold value
367   was applied to the cloglog output, calculating the effective cutoff as a mean of all model
368   repetitions threshold values, corresponding to a cutoff value of 0.0644, the areas below which
369   were then considered as AMF non-idoneity zones and plotted in black on the resulting map (Fig.
370   8A). A mean of all model repetitions values of 10 percent fixed cumulative thresholds
371   transformed into cloglog probability values generated by MaxEnt was chosen as the effective
372   cutoff for the cumulative output, corresponding to a value of 0.1266, the areas below which were
373   also in this case considered as AMF non-idoneity zones and plotted in black on the cumulative
374   output map (Fig. 8B). Actual absence points (resulting from the OTU table analysis) were not
375   used in this process due to a non-systematic coverage of the predicted territory, thus preferring
376   not to introduce biases in the MaxEnt only-presence environmental suitability predictions.
377   In a further set of models excluding Chaupi Rodeo points of presence, the values of importance
378   of predictors were consistent with the full points model, except for elevation and type of
379   vegetation. In these models, elevation exhibited negligible importance and the type of vegetation
380   was uninformative, because the used points of presence were all included into the
381   phytogeographic zone of Puna. Temperature annual range (bio7) and soil type appeared to be the
382   best predictors also in this new set of models (Table S6). The output maps generated excluding
383   Chaupi Rodeo points evidenced a lower AMF suitability of Chaupi Rodeo area relative to the
384   full point outputs, suggesting that the agricultural land use may be a main driver of AMF
385   distribution in this area. The threshold types chosen as presence-absence cutoffs were the same
386   selected for the full points models respectively with values of 0.0474 and 0.1143 for the cloglog

387   and the cumulative models (Fig. 9A,B). According to the response curves of the generated
388   MaxEnt models, the selection of climatic predictors indicates a positive effect of temperature
389   seasonality in shaping the AMF Argentinean Puna distribution, while a higher altitude is
390   negatively correlated with the presence of AMF (Fig. 10).
391   According to the response curves of the generated MaxEnt models, the selection of climate
392   predictors indicates a positive effect of temperature seasonality in modeling the distribution of
393   the Argentine Puna AMF, while a higher altitude is negatively correlated with the presence of
394   AMF (Fig. 10).

395   **DISCUSSION**

396   Environmental metabarcoding outcomes indicated a high occurrence of Ascomycota and
397   Basidiomycota in the investigated Argentinean Puna and Altoandino soils, as expected given the
398   wide taxonomic diversity of these Phyla, and of the Phylum Mortierellomycota with lower
399   number of reads. Even if other studies assessed a widespread dominance of Chytrids in high-
400   elevation periglacial soils (Freeman et al., 2009) we found very scarce occurrence of these taxa
401   in Puna soil, likely due to an endemic aridity in most of the year which may not favor the Chytrid
402   zoosporic reproductive stages. The Phylum Glomeromycota occurred at higher abundances in
403   croplands than in uncultivated Puna and Salar soils, likely due to a greater density of host plants
404   in the former ones relative to the typical scarce vegetation of the habitats of these areas.
405   Enhanced mycorrhizal dynamics and nutrients exchange certainly contributed to shaping
406   different physical and biochemical features in Chaupi Rodeo farmed soils, which are not
407   intensively managed and likely turn out to be more fit to Glomeromycota growth than droughty
408   and poorly evolved soils of Puna environments. Still, this is not a univocal trend, since potato
409   crop fields showed a very low occurrence of AMF not only relative to corn and fava beans but
410   also in comparison to other Puna soils, AMF being virtually absent in two potato crop samples
411   out of three. In a previous study focused on the Chaupi Rodeo cropland soils, Ontivero et al.,
412   (2020) underlined how AMF communities were significantly shaped by calcium and nitrogen
413   concentration in soils. Similarly, other studies highlighted how pH and calcium concentration
414   was strongly correlated with fungal richness in soils, with particular significance of pH and
415   evapotranspiration for Glomeromycota occurrence (Tedersoo et al., 2014). In a SDM research on
416   the realized niche of AMF, Davison et al., (2021) showed how pH and temperature were the
417   most significant drivers of the global AMF distribution, reinforcing the importance of climatic
418   and biogeochemical features of soils for explaining the AMF distribution.
419   Among all locations of Puna environments, Dunas samples and to follow both Puesto del
420   Marqués and Punto Susques samples showed the lowest occurrence of AMF. If the higher
421   salinity of Dunas soil might additionally contribute to lower the AMF presence, for all the above
422   mentioned locations the AMF scarcity is likely related to soil disturbance due to grazing, flocks
423   of llamas and cattle being restricted in paddocks or in corrals in Puesto del Marqués, and only
424   llamas in Dunas and Punto Susques. The effects of grazing on AMF are however controversial
425   because the responses of these biotrophic symbiotic fungi to herbivory are context-dependent
426   and may be directly related to the carbon flux within the plant-AMF-soil network and to the
427   intensity and extent of grazing over time, as well as to the specific mycorrhizal dependence of
428   the given grazed plants and the adaptation of AMF and their host plants to grazing (van der
429   Heyde et al., 2019). Modeling of AMF responses to grazing by considering soil hyphal length,
430   host root colonization, composition of soil communities and their sporulation, highlighted that
431   these parameters were driven by the extent of grazing time (van der Heyde et al., 2017) in

432 Canadian grasslands with 17 to 85 years of grazing by cattle and horses. In Puna, anthropogenic
433 activity since the late Holocene, and for more than 2,000 years before the present, led to intense
434 grazing, which is a major cause of shrubland expansion to the detriment of grassland habitats.
435 Furthermore, grazing by domestic animals including cows and camelids proved to be a main
436 driver of more drastic effects on vegetation than those caused only by wild camelid species
437 (Carilla et al., 2018, Quiroga Mendiola & Cladera 2018), thus impacting on soil fungal dynamics
438 and likely on AMF distribution as well. Overall, a global meta-analysis of livestock effect on
439 AMF has shown the negative effect of heavy or moderate grazing on AMF abundance, as well as
440 for longtime grazing and low annual precipitations, consequently triggering the reduction of host
441 aboveground biomass (Yang et al., 2020), as recorded modeling the sequences of Puna. The
442 relationship between AMF and soil total fungal profile with grazing-shaped vegetation dynamics
443 in Puna ecosystems requires, however, further investigation.
444   The occurrence of AMF in hypersaline Salar areas and surroundings, notably the locations of
445 Salinas Grandes and Dunas is very scarce. No AMF sequences were recorded in any of the
446 Salinas Grandes soils, while only one sampling point in Dunas displayed the occurrence of a
447 single Diversisporaceae OTU. As the model averaged estimates statistically suggest, it is
448 therefore likely that high levels of salinity in soils can have a negative impact on the overall
449 AMF occurrence and abundance in these environments, this effect extending also to areas in
450 moderately saline surroundings of the Salar. It must be noted however that Glomeromycota are
451 obligate biotrophic symbionts of most vascular plants and the observed variability in abundance
452 and distribution of sequences from these fungi among different points of the same locations
453 might be partly driven by the chance that sampling occurred at different distances from nearby
454 plants or plant root residuals and to plant-host specificity at the ecological group level (Opik et
455 al., 2009). However, results obtained so far in AMF morphological studies in saline
456 environments of Argentina did not show a clear pattern with regard to the effect of plant identity
457 on AMF occurrence and abundance. Different native halophytes of Salinas of central Argentina
458 showed in their rhizosphere a low diverse AMF community, with AMF species inconsistently
459 varying in their sporulation among soil samples, seasons, plant species and soil depth (Becerra et
460 al., 2014; Cofré et al., 2012; Soteras et al., 2012). Therefore, the effect of salinity on AMF
461 occurrence is not straightforward.
462 Along with salinity, the abundance of organic matter in soil is suggested to be a significant
463 predictor of AMF distribution in the investigated soils, in accordance with the importance of soil
464 type and nutrient abundance and retention highlighted by MaxEnt models. Even if
465 Glomeromycota are not saprotrophs, a larger amount of organic matter in soil may be an
466 indicator of enhanced nutrient soil dynamics generated by a greater density of plants, thus once
467 more likely explaining the abundance of AMF in crop soils compared to soils of Puna and Salar,
468 typically poor in organic matter. MaxEnt response curves regarding edaphic and vegetation
469 variables (Fig.10, Fig. S4-8) pointed out that grasslands, followed by sparse herbaceous and
470 shrubs areas, are more suitable for the occurrence of AMF than bare areas, as expected for a
471 taxon in obligate symbiosis with plant roots. Soil types predominantly associated with a greater
472 occurrence of AMF in Puna environments are luvic yermosols, with a weak ochric A horizon and
473 an argillic B horizon, aridic moisture regime and low organic carbon content, mainly
474 corresponding to soil characteristics of Puna habitats. In a recent research Větrovský et al.,
475 (2019) assessed that climate is the primary environmental factor for the overall distribution of
476 Fungi. Specifically, temperature proved to be a main driver for AMF distribution, as shown by
477 Zhao et al., (2019) and Davison et al., (2021).

478 In Argentinean Puna sites, MaxEnt response curves also suggested that a wider annual
479 temperature range showed a greater positive effect on the AMF distribution among the other
480 bioclimatic predictors of AMF seasonality, while higher altitude was negatively correlated with
481 AMF presence. The fact that an increasing temperature seasonality can be related to a greater
482 occurrence of AMF might appear counter-intuitive because it suggests more extreme climatic
483 conditions. The effect of temperature on AMF is related directly or indirectly to carbon and
484 phosphorus AMF-host exchange and to translocations from soil to the host plants (Gavito et al.,
485 2005). Recent research highlighted that AMF showed differential patterns of growth depending
486 on the environmental conditions of their habitats, with a higher resistance to high temperature in
487 arid and semiarid ecosystems (Kilpeläinen et al., 2020). Therefore, in the ecological context of
488 the Puna environment, it can be inferred that in the wet and warm season AMF can benefit from
489 higher temperatures, enhanced water availability and metabolic activity, likely remaining
490 dormant as spores in less favorable periods. The seasonal distribution of AMF is well known for
491 their sporulation (Smith & Read 2008), measured in terms of spore abundance and sporulating
492 species richness and root colonization, especially for other investigated South American highland
493 grasslands such as the Pampa de Achala (Lugo & Cabello, 2002; Lugo et al., 2003) and different
494 mountain environments (Soteras et al., 2019). The altitude negative trend might be related to less
495 favorable environmental and climatic conditions at higher altitudes for plants and therefore for
496 AMF, as it has been reported in Argentinean Puna considering sporulation (Lugo et al., 2008)
497 and root colonization (Lugo et al., 2012). A study performed in the Andean Yungas Forest
498 ecoregion (Geml et al., 2014) assessed how soil-fungal community structure is negatively
499 correlated with elevation, in accordance with former studies on AMF distribution on Himalaya
500 (Gai et al., 2012) and our results on AMF sequences distribution. Although a wider sampling in
501 high elevation gramineous grasslands might be required to better explain the AMF ecology in
502 these habitats, AMF community richness have already been assessed to be higher in grasslands at
503 global (Tedersoo et al., 2014, Davison et al., 2015) and local scales (Grilli et al., 2019).
504 Given the evidence that croplands hosted a much greater abundance of Glomeromycota, we
505 considered it advisable to account for the human intervention in cultivated soils in the modeling
506 process of AMF distribution. Setting a new series of models excluding the cultivated Chaupi
507 Rodeo samples allowed us to model the potential AMF distribution as if these areas were not
508 cultivated, verifying how large-scale soil features and bioclimatic constraints would act on the
509 AMF distribution if these soils had remained unmanaged habitats of Puna and Altoandino. Such
510 a new set of models unveiled a much lower suitability of Chaupi Rodeo location for
511 Glomeromycota than previously modeled, but also at a wider scale reshaped the distribution of
512 AMF in the areas where climatic and soil type conditions were similar to Chaupi Rodeo
513 locations. It is noteworthy that the new models excluded lithosols from high suitability: these
514 soils are typically found in the unfarmed Chaupi Rodeo areas, but in the previous models they
515 were inversely considered as good predictors of AMF occurrence (see Fig. S4-6). This confirms
516 that shaping distribution models without accounting for human activity in the studied
517 environment may prove misleading but likewise it allows to model potential distributions more
518 accurately in a wider ecological perspective.

CONCLUSIONS

520 To sum up, we performed an exploratory assessment of AMF environmental suitability in an
521 ecologically homogeneous area of Argentinean Puna and Altoandino throughout three different
522 representative ecosystems: punean grazed and ungrazed grasslands, a hypersaline endorheic

523 basin and family-farmed croplands. We concluded that the differential human impact on frail
524 ecosystems like the Argentinean Puna should be assessed also for soil and microbiological
525 dynamics, especially when investigating deep plant-related organisms like AMF so essential in
526 preserving not only the ecosystems stability but also the farming productivity, this last
527 unreplaceable in the subsistence-based human communities of Puna. Likewise, AMF distribution
528 may be a good indicator of derangements from ecosystem equilibrium in those locations of Puna
529 where exceeding intensive grazing, wood collection and mining activities risk jeopardizing the
530 good functioning of these unique ecosystems. We encourage further investigations of the
531 structure and composition of punean soil mycobiota at a larger scale and with an in-depth
532 assessment of the differential impact of human activities on the many habitats composing this
533 nowadays endangered biogeographical region.

534

Competing Interests

Author Contributions

537 **Davide Nepote Valentin** and **Samuele Voyron** conceived and designed the analyses of data,
538 search strategy and analyzed data, prepared figures and/or tables, authored or reviewed drafts of
539 the paper and approved the final draft.
540
541 **Florencia Soteras** discussed strategy, analyzed data, authored or reviewed drafts of the paper,
542 approved the final draft
543
544 **Hebe J. Iriarte** prepared figures and/or tables, analyzed data, authored or reviewed drafts of the
545 paper, and approved the final draft.
546
547 **Andrea Giovannini** supervised the analysis of data and discussed strategy, authored or reviewed
548 drafts of the paper, and approved the final draft.
549
550 **Erica Lumini** and **Mónica A. Lugo** conceived and designed the experiments and the field
551 campaign, performed the sampling and the experiments, supervised the analysis of the data,
552 contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, and
553 approved the final draft.
554
555

Data Availability

557 The following information was supplied regarding data availability:
558 The raw measurements are provided in the Supplementary Files.
559
560

561 **REFERENCES**

562 **Abarenkov K, Zirk A, Piirmann T, Pöhönen R, Ivanov F, Nilsson RH, Kõljalg U.** 2021.
563 UNITE QIIME release for Fungi 2. Version 10.05.2021. UNITE Community.
564 https://doi.org/10.15156/BIO/1264763

565 **Austin MP.** 2002. Spatial prediction of species distribution: an interface between ecological
566 theory and statistical modelling. *Ecological Modelling* **157**:101/118.

567 **Barton K. 2020.** Mu-MIn: Multi-model inference. R Package. Version 1.43.17. http://R-
568 Forge.R-project.org/projects/mumin/

569 **Bradie J & Leung B.** 2017. A quantitative synthesis of the importance of variables used in
570 MaxEnt species distribution models. *Journal of Biogeography* **44**:1344–1361.

571 **Becerra A, Bartoloni N, Cofré N, Soteras F, Cabello M.** 2014. Arbuscular mycorrhizal fungi
572 in saline soils: Vertical distribution at different soil depth. *Brazilian Journal of Microbiology*
573 **45(2)**:585–594.

574 **Berruti A, Desirò A, Visentin S, Zecca O, Bonfante P.** 2017. ITS fungal barcoding primers
575 versus 18S AMF-specific primers reveal similar AMF-based diversity patterns in roots and soils
576 of three mountain vineyards. *Environmental Microbiology* **9**:658–667.

577 **Bolyen E, Rideout JR, Dillon MR et al.** 2019. Reproducible, interactive, scalable and
578 extensible microbiome data science using QIIME 2. *Nature Biotechnology* **37**:852–857.

579 **Bonfante P & Genre A. 2015.** 2015. Arbuscular mycorrhizal dialogues: do you speak 'plantish '
580 or 'fungish'? *Trends in Plant Science* **20**:150–154.

581 **Cabrera AL.** 1976. Territorios fitogeográficos de la República Argentina. Pp. 1-85 in W. F.
582 Kugler (ed.). *Enciclopedia Argentina de Agricultura y Jardinería* **2**. Acme, Buenos Aires,
583 Argentina, Fascículo 1.

584 **Cabrera AL & Willink A.** 1980. Biogeografía de América Latina. *OEA*, Washington, DC.

585 **Callahan BJ, MCMurdie PJ, Rosen MJ et al.** 2016. DADA2: high-resolution sample inference
586 from Illumina amplicon data. *Nature Methods* **13**:581.

587 **Carilla J, Grau A., Cuello S.** 2018. Vegetación de la Puna argentina. In: *Serie Conservación de
588 la Naturaleza 24*: *La Puna argentina: naturaleza y cultura* **7**.

589 **Chong J, Liu P, Zhou G, Xia J.** 2020. Using MicrobiomeAnalyst for comprehensive statistical,
590 functional, and meta-analysis of microbiome data. *Nature Protocols* **15**:799–821
591 DOI: 10.1038/s41596-019-0264-1.

592 **Cofré N, Becerra A, Nohura E, Soteras F.** 2012. Arbuscular mycorrhizae and dark-septate
593 endophytes on Atriplex cordobensis in saline sites from Argentina. *Journal of Agricultural
594 Technology* **8**:2201–2214.

595 **Davison J, Moora M, Öpik M, Adholeya A, Ainsaar L, Bâ A, Burla S, Diedhiou AG,
596 Hiiesalu I, Jairus T, Johnson NC, Kane A, Koorem K, Kochar M, Ndiaye C, Pärtel M,
597 Reier Ü, Saks Ü, Singh R, Vasar M, Zobel M.** 2015. Global assessment of arbuscular
598 mycorrhizal fungus diversity reveals very low endemism. *Science* **127(6251)**:970-973.

599 **Davison J, Moora M, Semchenko M et al.** 2021. Temperature and pH define the realised niche
600 space of arbuscular mycorrhizal fungi. *New Phytologist* **231**:763–776.

601 **Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J.** 2017. MicrobiomeAnalyst: a
602 web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data.
603 *Nucleic Acids Research* **45(W1)**:W180–W188. https://doi.org/10.1093/nar/gkx295.

604 **Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH.** 2010. Relative roles of niche and
605 neutral processes in structuring a soil microbial community. *The ISME Journal* **4**:337–345.

606 **Fox J & Weisberg S.** 1976. An R Companion to Applied Regression, Third edition. Sage,
607 Thousand Oaks CA.

608 **Freeman KR, Martin AP, Karki D et al.** 2009. Evidence that chytrids dominate fungal
609 communities in high-elevation soils. *PNAS* **106**:18315–20.

610 **Gai JP, Tian H, Yang FY, Christie P, Li XL, Klironomos JN.** 2012. Arbuscular mycorrhizal
611 fungal diversity along a Tibetan elevation gradient. *Pedobiologia* **55**:145–151.

612 **Gavito ME, Olsson PA, Rouhier H, Medina-Peñafiel A, Jakobsen I, Bago A, Azcón-Aguilar**
613 **C.** 2005. Temperature constraints on the growth and functioning of root organ cultures with
614 arbuscular mycorrhizal fungi. *New Phytologist* **168**:179–188. doi.org/10.1111/j.1469-
615 8137.2005.01481.x

616 **Grilli G, Marro N, Risio Allione L.** 2019. Structure and diversity of arbuscular mycorrhizal
617 fungal communities across spatial and environmental gradients in the Chaco Forest of South
618 America. In: *Mycorrhizal Fungi in South America*, Springer.

619 **Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ.** 2011. A statistical explanation of
620 MaxEnt for ecologists. *Diversity and Distributions* **17**:43–57.

621 **Fick SE, Hijmans RJ.** 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for
622 global land areas. International Journal of Climatology **37(12)**:4302-4315.
623 https://doi.org/10.1002/joc.5086

624 **Fox J & Weisberg S.** 2019. An R Companion to Applied Regression, Third edition. Sage,
625 Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

626 **Geml J, Pastor N, Fernandez L, Pacheco S, Semenova TA, Becerra AG, Wicaksono CY,**
627 **Nouhra ER.** 2014. Large-scale fungal diversity assessment in the Andean Yungas forests reveals
628 strong community turnover among forest types along an altitudinal gradient. *Molecular Ecology*
629 **23**: 2452–2472.

630 **Hammer Ø, Harper DAT, Paul DR.** 2001. Past: Paleontological Statistics Software Package
631 for Education and Data Analysis. *Palaeontologia Electronica* **4(1)**: art. 4.

632 **Hijmans RJ, Phillips SJ, Leathwick J, Elith J.** 2021. Dismo: Species Distribution Modeling. R
633 Package Version 1.1.3-5. http://CRAN.R-project.org/package=dismo.

634 **Kilpeläinen J, Aphalo PJ, Lehto T.** 2020. Temperature affected the formation of arbuscular
635 mycorrhizas and ectomycorrhizas in Populus angustifolia seedlings more than a mild drought.
636 *Soil Biology & Biochemistry* **146**:107798. https://doi.org/10.1016/j.soilbio.2020.107798

637 **Korner-Nievergelt F, Roth T, von Felten S, Guelat J, Almasi B, Korner-Nievergelt P.** 2015.
638 Bayesian Data Analysis in Ecology using Linear Models with R, BUGS and Stan. *Elsevier*.

639 **Lugo MA & Cabello MN.** 2002. Native arbuscular mycorrhizal fungi (AMF) from mountain
640 grassland (Córdoba, Argentina) I. Seasonal variation of fungal spore diversity. *Mycologia* **94(4)**:
641 579–586.

642 **Lugo MA, González Maza ME, Cabello MN.** 2003. Arbuscular mycorrhizal fungi (AMF) from
643 native Argentinian -Southamerican- mountain grassland II. Seasonal variation of colonization
644 and its relation with grazing and metabolic host type. *Mycologia* **95(3)**:407–415.

645 **Lugo MA, Ferrero M, Menoyo E et al.** 2008. Arbuscular mycorrhizal fungi and rhizospheric
646 bacteria diversity along an altitudinal gradient in South American Puna grassland. *Microbial*
647 *Ecology* **55**:705–713.

648 **Lugo MA & Menoyo E.** 2019. Chapter 12: Southern Highlands: Fungal Endosymbiotic
649 Associations. In: Mycorrhizal Fungi in South America; Life Science Series, Fungal Biology;
650 Pagano MC, Lugo MA, Eds.; Springer: Cham, Switzerland, 217–255.

651 **Lugo MA, Negritto A, Jofré M, Anton A, Galetto L.** 2012. Colonization of native Andean
652 grasses by arbuscular mycorrhizal fungi: a matter of altitude, host photosynthetic pathway and
653 host life form? *FEMS Microbial Ecology* **81**: 455–466. DOI: 10.1111/j.1574-
654 6941.2012.013732012.x.

655 **Martínez Carretero EM.** 1995. La Puna argentina. Delimitación general y división en distritos
656 florísticos. *Bol.Soc.Argent.Bot*. **31(1-2)**:27-40.

657 **McMurdie PJ & Holmes S.** 2013. phyloseq: an R package for reproducible interactive analysis
658 and graphics of microbiome census data. *Plos One* **8**:e61217.

659 **Matteucci SD.** Chapter 2: Ecorregión Altos Andes. In: Morello J, Matteucci SD, Rodríguez AF,
660 Silva M (eds.), Ecorregiones y Complejos ecosistémicos argentinos, 2nd ed. Facultad de
661 Arquitectura, Diseño y Urbanismo. Grupo de Ecología del Paisaje y Medio Ambiente.
662 Universidad de Buenos Aires. GEPAMA 2018 a: 17- 108 pp.

663 **Matteucci SD.** Chapter 3: Ecorregión Puna. In: Morello J, Matteucci SD, Rodríguez AF, Silva
664 M (eds.), Ecorregiones y Complejos ecosistémicos argentinos, 2nd ed. Facultad de Arquitectura,
665 Diseño y Urbanismo. Grupo de Ecología del Paisaje y Medio Ambiente. Universidad de Buenos
666 Aires. GEPAMA 2018 b:109- 152 pp.

667 **Morrone JJ.** 2001. Biogeografía de América Latina y el Caribe. *M&T–Manuales & Tesis SEA*,
668 **vol. 3**. Zaragoza.

669 **Nguyen NH, Song Z, Bates ST et al.** 2016. FUNGuild: An open annotation tool for parsing
670 fungal community datasets by ecological guild. *Fungal Ecology* **20**:241–48.

671 **Ondov BD, Bergman NH, Phillippy AM.** 2011. Interactive metagenomic visualization in a
672 Web browser. *BMS Bioinformatics* **12**:385.

673 **Ontivero RE, Voyron S, Risio Allione LV, Bianco P, Bianciotto V, Iriarte HJ, Lugo MA,**
674 **Lumini E.** 2020. Impact of land use history on the arbuscular mycorrhizal fungal diversity in
675 arid soils of Argentinean farming fields. *FEMS Microbiology Letters* **367(14)**: fnaa114.
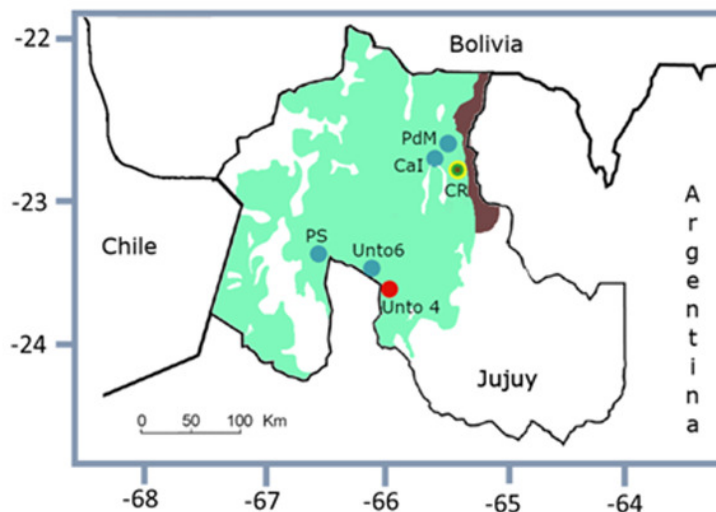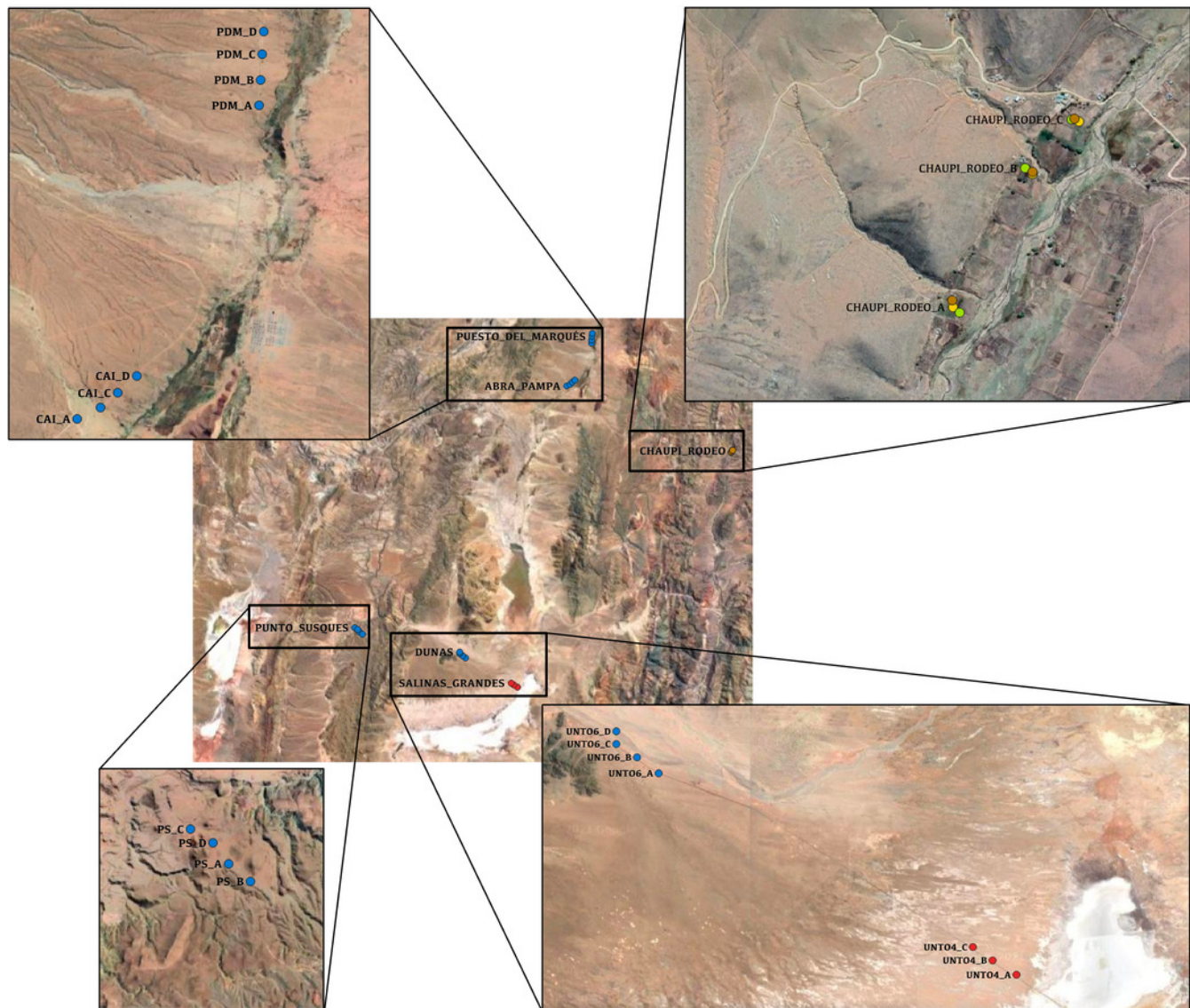
676 **Öpik M , Metsis M, Daniell T J, Zobel M, Moora M.** 2009. Large-scale parallel 454
677 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a
678 boreonemoral forest. *New Phytologist* **184(2)**:424-437.

679 **Oyarzabal M, Clavijo J, Oakley L, Biganzoli F, Tognetti P, Barberis I, Maturo HM,**
680 **Aragon R, Campanello PI, Prado D, Oesterheld M, Leon RJC.** 2018. Unidades de vegetación
681 de la Argentina. *Ecología Austral* **28**:040-063.

682 **Panigatti, J.L.** 2010. Argentina: 200 Años, 200 Suelos; INTA: Buenos Aires, Argentina; p. 345.
683 https://inta.gob.ar/sites/default/files/script-tmp-inta-200-suelos.pdf

684 **Phillips SJ, Anderson RP, Schapire RE.** 2006. Maximum entropy modeling of species
685 geographic distributions. *Ecological Modelling* **190**:231-259.

686 **Phillips SJ, Dudik M, Schapire RE.** 2004. A maximum entropy approach to species
687 distribution modeling. Pages 655-662 in Proceedings of the 21st International Conference on
688 Machine Learning. ACM Press, New York.

689 **QGIS.org.** QGIS Geographic Information System. QGIS Association 2006. http://www.qgis.org

690 **Quiroga Mendiola M, Cladera J.** 2018. Ganadería en la Puna argentina. In: HR Grau, J Babot,
691 A Izquierdo, A Grau (eds.). La Puna Argentina: naturaleza y cultura. *Serie Conservación de la*
692 *Naturaleza* **24**:387-402.

693 **R Core Team.** 2013. R: A language and environment for statistical computing. R Foundation for
694 Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

695 **Renison D, Cuyckens Gae, Pacheco S, Guzmán G, Grau H, Marcora P, Robledo G,**
696 **Cingolani Am, Dominguez J, Landi M, Bellis L, Hensen I.** 2013. Distribución y estado de
697 conservación de las poblaciones de árboles y arbustos del género Polylepis (Rosaceae) en las
698 montañas de Argentina. *Ecología Austral* **23**: 27-36.

699 **Soteras F, Becerra A, Cofré N, Bartoloni, Cabello M.** 2012. Arbuscular mycorrhizal fungal
700 species in saline environments of Central Argentina: seasonal variation and distribution of spores
701 at different soil depth. *Sydowia* **64**:301–311.

702 **Soteras F, Menoyo E, Grilli G et al.** 2019. Arbuscular mycorrhizal fungal communities of high
703 mountain ecosystems of South America: relationship with microscale and macroscale factors. In:
704 Pagano MC, Lugo MA (eds). Mycorrhizal Fungi in South America. Switzerland: Springer, 257–
705 275.

706 **Tedersoo L, Bahram M, Polme S et al.** 2014. Global diversity and geography of soil fungi.
707 *Science* **346**:1256688.

708 **Troll C.** 1959. Die Tropischen Gebirge. Ihre dreidimensionale klimatische und
709 pflanzengeographische zonierung. *Bonner Geogr. Abh. Heft* **25**:23 p.

710 **Troll C.** 1968. The Cordilleras of the Tropical Americas. Aspects of climatic, phytogeographical
711 and agrarian ecology. Geo-Ecology of the Mountains Regions of the Tropical Americas; 15–56.

712 **Turner JCM & Méndez V.** 1979. Puna, en: Geología Regional Argentina 2°Simposio.
713 *Acad.Nac.Ci.* **1**:13-56.

714 **van der Heyde M, Abbott LK, Gehring C, Kokkoris V, Hart MM.** 2019. Reconciling
715 disparate responses to grazing in the arbuscular mycorrhizal symbiosis. *Rhizosphere* **11**:1–7,
716 doi.org/10.1016/j.rhisph.2019.100167.

717 **van der Heyde M, Bennett JA, Piter J, Hart MM.** 2017. Longterm effects of grazing on the
718 arbuscular mycorrhizal symbiosis. Agriculture, Ecosystems & Environment **243**: 27–33.
719 http://dx.doi.org/10.1016/j.agee.2017.04.003

720 **Větrovský T, Kohout P, Kopecký M et al.** 2019. A meta-analysis of global fungal distribution
721 reveals climate-driven patterns. *Nature Communications* **10**:5142.

722 **Vorano A & Vargas Gil R.** 2002. Evaluación de la situación actual de los procesos de
723 Desertificación de la Puna Salto Jujeña. Soluciones alternativas. INTA Estación Experimental
724 Agropecuaria Salta, Salta.
725

726 **Warren DL & Seifert SN.** 2011. Ecological niche modeling in Maxent: the importance of
727 model complexity and the performance of model selection criteria. *Ecological Applications*
728 **21(2)**:335–342.

729 **Wei T & Simko V.** 2021. R package 'corrplot': Visualization of a Correlation Matrix. (Version
730 0.92). https://github.com/taiyun/corrplot.

731 **Wickham H.** 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
732 ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

733 **Wijayawardene NN, Hyde KD, Al-Ani LKT et al.** 2020. Outline of fungi and fungus-like taxa.
734 Mycosphere **11**:1060–456.

735 **Wurzburger N, Brookshire ENJ, McCormack ML, Lankau RA.** 2017. Mycorrhizal fungi as
736 drivers and modulators of terrestrial ecosystem processes. *New Phytologist* **213**:996–999.

737 **Yang X, Chen J, Shen Y, Dong F, Chen J.** 2020. Global negative effects of livestock grazing
738 on arbuscular mycorrhizas: A meta-analysis. *The Science of the total environment* **708**:134553.
739 doi.org/10.1016/j.scitotenv.2019.134553

740 **Zhao X, Yang Y, Shen H, Geng X, Fang, J.** 2019. Global soil–climate–biome diagram: linking
741 surface soil properties to climate and biota. *Biogeosciences* **16**:2857–2871.

742 **Zimmermann NE, Edwards TC Jr., Graham CH, Pearman PB, Svenning JC.** 2010. New
743 trends in species distribution modelling. *Ecography* **33**:985-989.

Figure 1

Geolocation of the analyzed sampling points within the Argentinean Jujuy province.

A. Points are clustered in locations labeled as follows: Unto4 (Salinas Grandes), Unto6 (Dunas), PS (Punto Susques), CaI (Camino al INTA Abra Pampa), PdM (Puesto del Marqués, Abra Pampa), CR (Chaupi Rodeo). **B**. Sampling points are shown. Chaupi rodeo points are labeled following the three sampling locations (A, B, C). Color legend shows different habitats or crops: red=salar, blue= puna, green= fava bean crops, yellow=corn crops, brown=potato fields.

Sampling points in Jujuy (Argentina):

- 🔴 Salar
- 🔵 Puna
- 🟡 CR

Crops in Chaupi Rodeo:

- 🟢 fava beans
- 🟡 corn
- 🟤 potato

Figure 2

Georeferenced, clipped and aligned rasters of the selected environmental predictors used for MaxEnt modeling along with the 19 Worldclim derived climatic variables.

A. Land Cover 30" (500m) (Global Land Cover by National Mapping Organizations - GLCNMO based on index LCCS developed by FAO); **B**. Vegetational units (Oyarzabal et al. 2018): Puna (Vegetation Unit 35), dominated by Fabiana densa and Baccharis boliviensis (light green) and Altoandino (Vegetation Unit 36), dominated by Senecio algens and Oxalys compacta (dark green). **C**. Soil nutrients retention, FAO Geonetwork. **D**. Soil nutrients availability, FAO Geonetwork. **E**. Soil type, DSMW-Digital Soil Map of the World (FAO); **F**. Elevation 30", Worldclim.
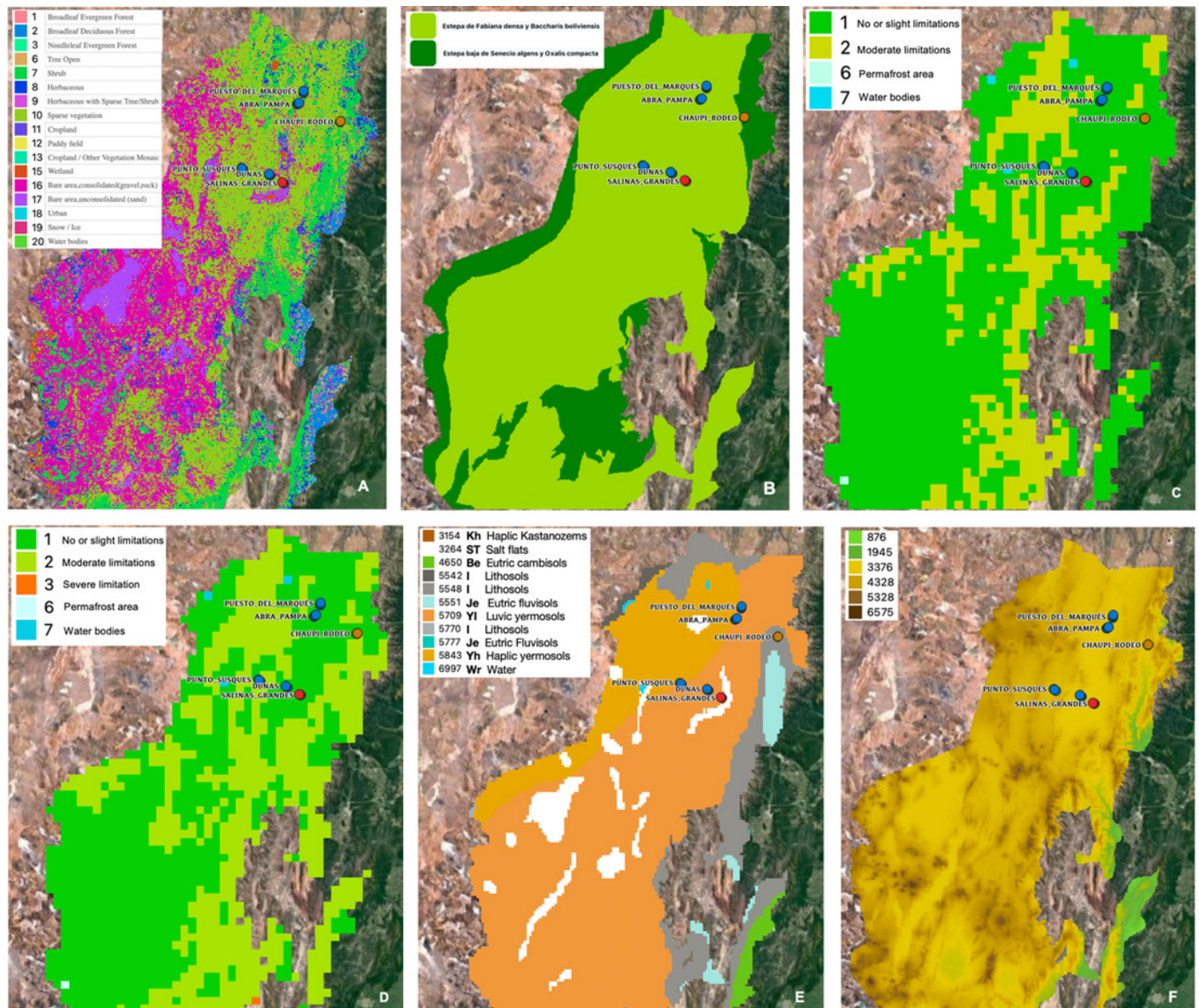
A
1	Broadleaf Evergreen Forest
2	Broadleaf Deciduous Forest
3	Needleleaf Evergreen Forest
6	Tree Open
7	Shrub
8	Herbaceous
9	Herbaceous with Sparse Tree/Shrub
10	Sparse vegetation
11	Cropland
12	Paddy field
13	Cropland / Other Vegetation Mosaic
15	Wetland
16	Bare area,consolidated(gravel,rock)
17	Bare area,unconsolidated (sand)
18	Urban
19	Snow / Ice
20	Water bodies

B
- Estepa de Fabiana densa y Baccharis boliviensis
- Estepa baja de Senecio algens y Oxalis compacta

C
1	No or slight limitations
2	Moderate limitations
6	Permafrost area
7	Water bodies

D
1	No or slight limitations
2	Moderate limitations
3	Severe limitation
6	Permafrost area
7	Water bodies

E
3154	Kh	Haplic Kastanozems
3264	ST	Salt flats
4650	Be	Eutric cambisols
5542	I	Lithosols
5548	I	Lithosols
5551	Je	Eutric fluvisols
5709	Yl	Luvic yermosols
5770	I	Lithosols
5777	Je	Eutric Fluvisols
5843	Yh	Haplic yermosols
6997	Wr	Water

F
876
1945
3376
4328
5328
6575

Figure 3

Taxonomic distribution of the retrieved fungal communities grouped per sample at the Phylum level.

Green=Ascomycota; Orange: Basidiomycota; Blue=unidentified fungi;

Purple=Glomeromycota; Light green=Mortierellomycota; Yellow=Chytridiomycota;
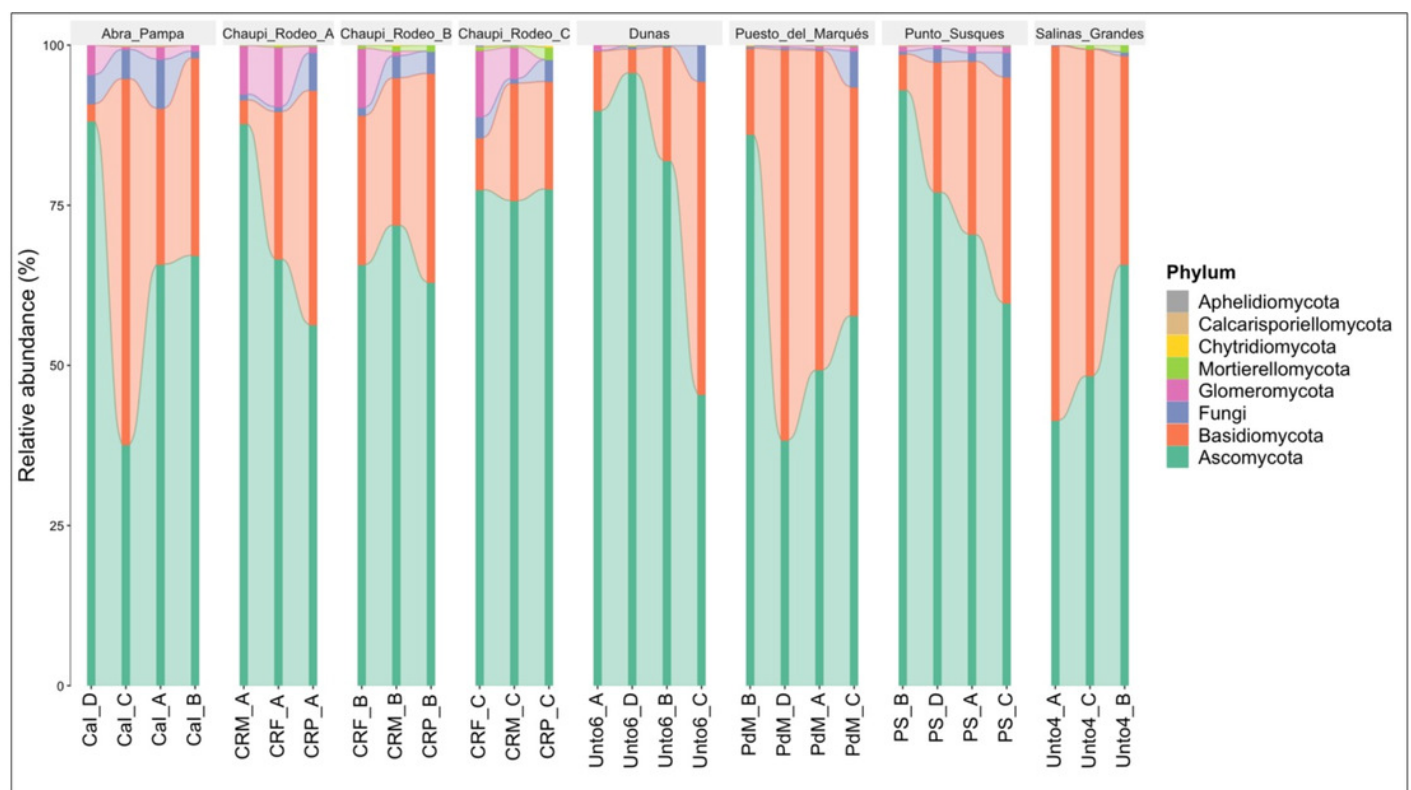
Brown=Calcarisporiellomycota; Grey=Aphelidiomycota.

Figure 4

Non-metric Multi-Dimensional Scaling (NMDS).

NMDS performed using Jaccard dissimilarity index and statistically tested by means of PERMANOVA, with a R-squared value of 0.41928 and p-value < 0.001.

Figure 5

Barchart of Glomeromycota read means among different habitats or land uses.

Samples habitats and land uses are indicated by colors as described in the legend. Different letters indicate pairwise post-hoc test statistical significances with Bonferroni correction (p<=0.05).

Figure 6

Training omission rate and predicted area performed in MaxEnt as a function of cumulative threshold.

Thresholds were averaged over the replicate runs respectively for Cloglog output (**A**) and for Cumulative output (**B**).

Figure 7

AMF environmental suitability as developed by MaxEnt iterative approach on the chosen best model.

A. Cloglog MaxEnt suitability map on a probability scale; **B**. Cumulative MaxEnt suitability map on a percent scale. Probabilities of occurrence are better represented by the cloglog output while habitat suitability is better represented by the cumulative output. Colors indicate a gradient from minimum suitability (deep blue) to maximum suitability (red).
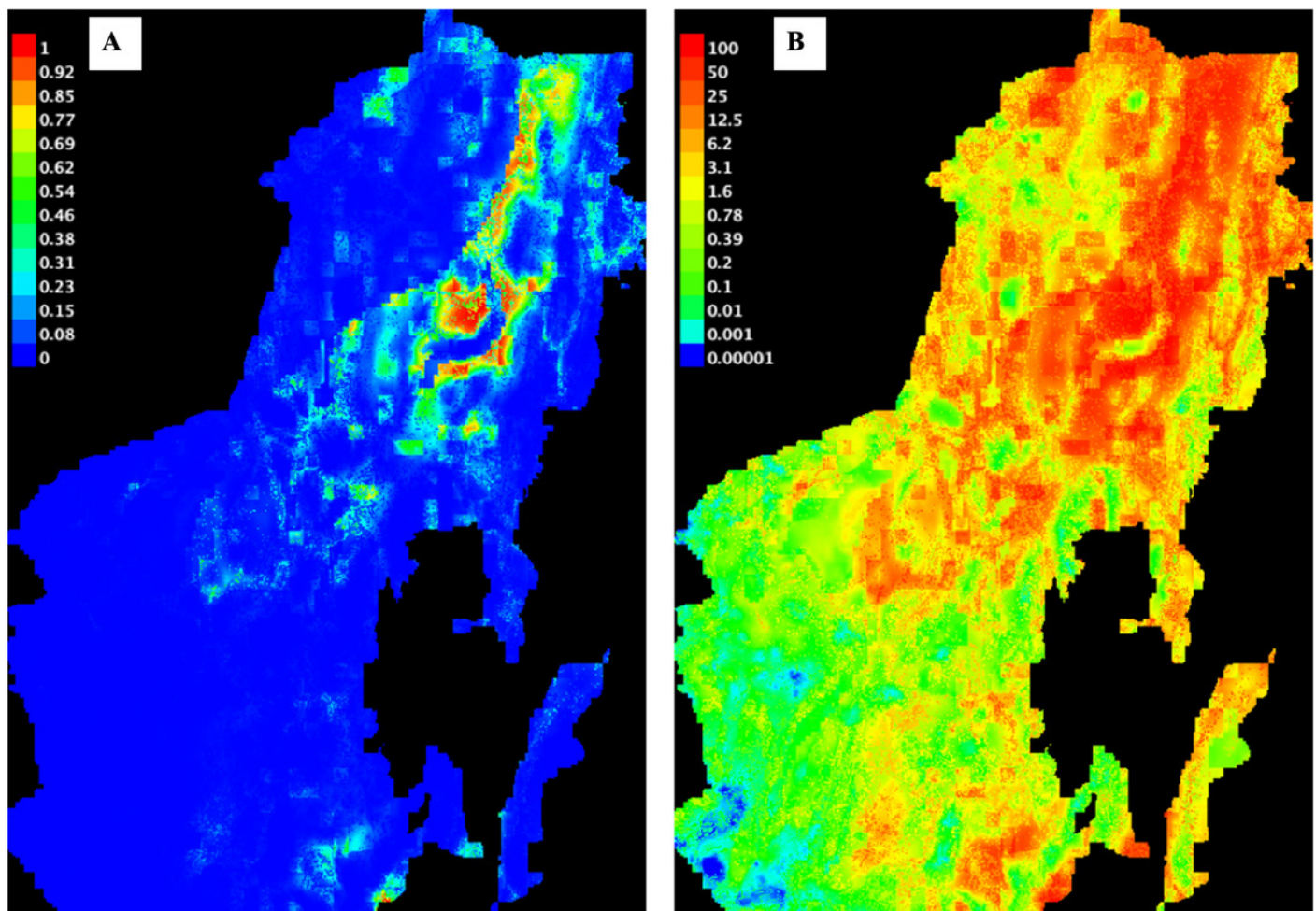
Figure 8

MaxEnt environmental suitability maps using all the points of presence.

A. MaxEnt cloglog model (threshold: 0.0644) **B**. Cumulative model) outputs (threshold: 0.1266). Black areas correspond to non-idoneity zones for AMF. Other colors represent a blue-red gradient from low to high environmental suitability for the occurrence of AMF. Presence points corresponding to the georeferenced sampling points where Glomeromycota sequences were found are plotted on the map as colored dots as follows: brown=farmed soils, green=puna, white=salar.
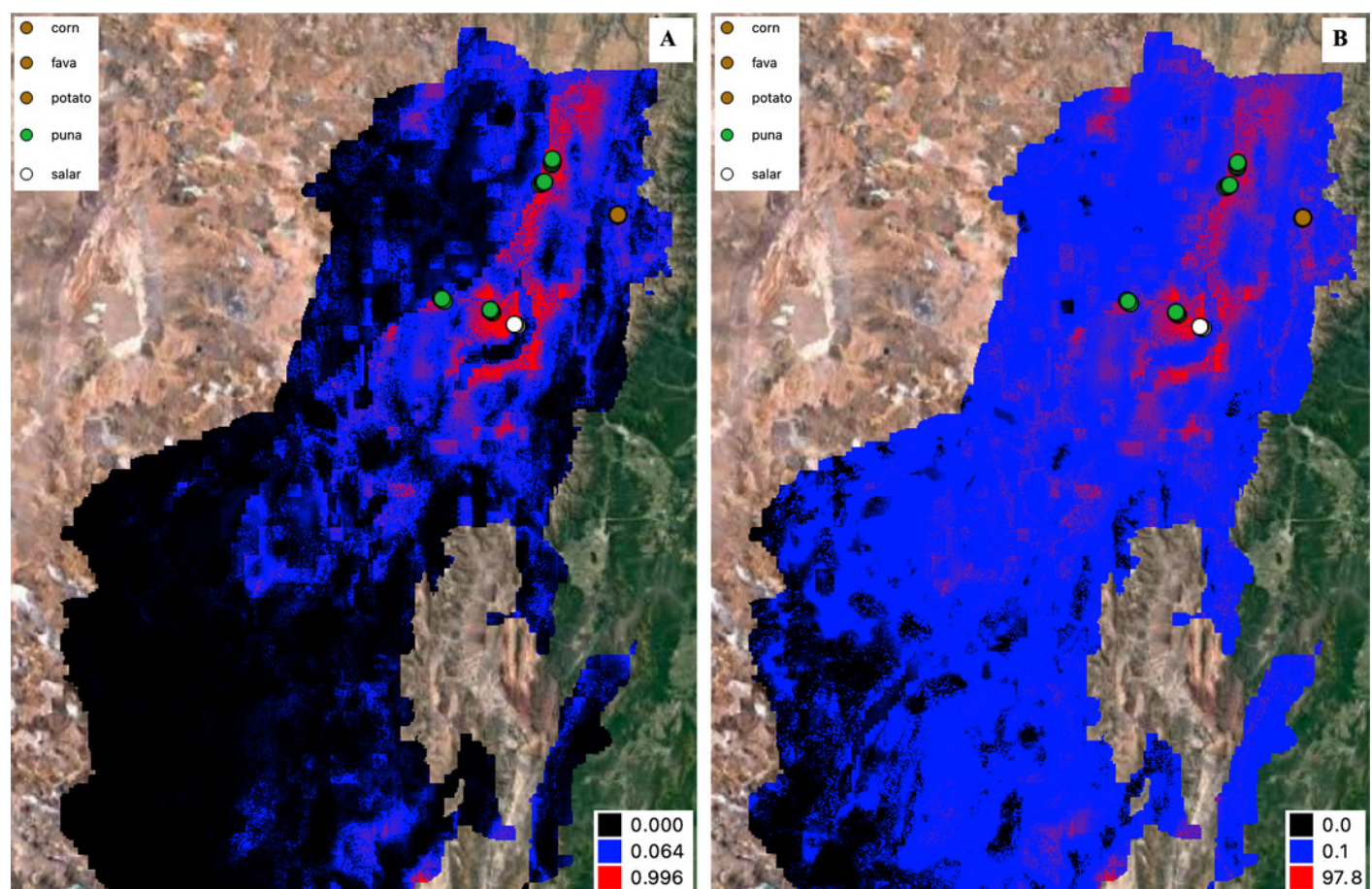
Figure 9

Environmental suitability maps generated using a subset of points of presence that exclude *Chaupi Rodeo* points.

A. MaxEnt cloglog model (threshold: 0.0474) **B**. Cumulative model (threshold: 01143). Black areas correspond to non-idoneity zones for AMF. Other colors represent a blue-red gradient from low to high suitability of occurrence of AMF. Other colors represent a blue-red gradient from low to high environmental suitability for the occurrence of AMF. Presence points corresponding to the georeferenced sampling points where Glomeromycota sequences were found are plotted on the map as colored dots as follows: brown=farmed soils, green=puna, white=salar. The points of *Chaupi Rodeo* are represented here for reference even if not included in MaxEnt models.
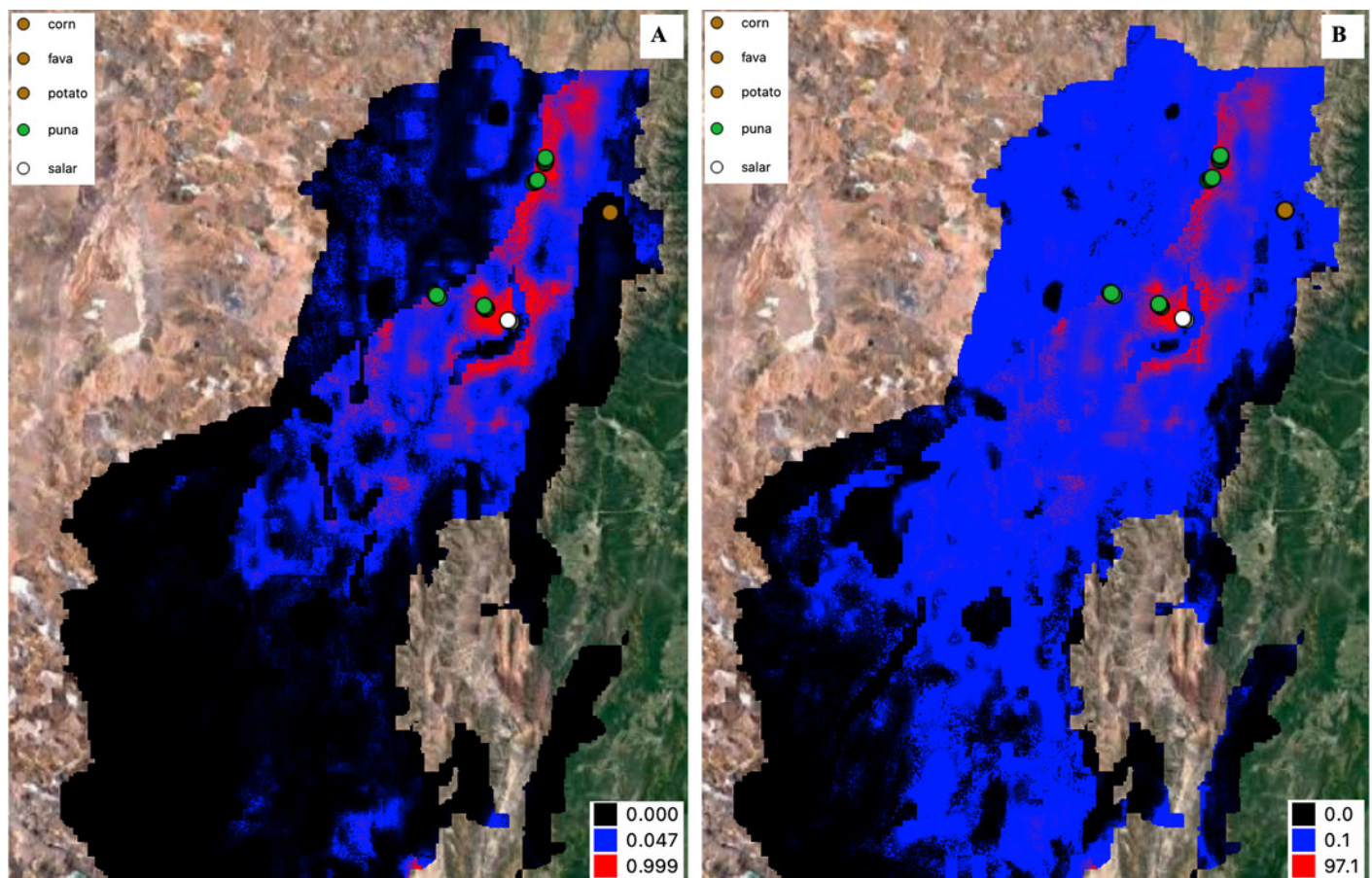
Figure 10

MaxEnt generated response curves for variables bio7, soil type and elevation extracted from the chosen cloglog full points model.

The curves show the effect of varying the chosen variable on the MaxEnt prediction, by keeping all other environmental variables at their average sample value.
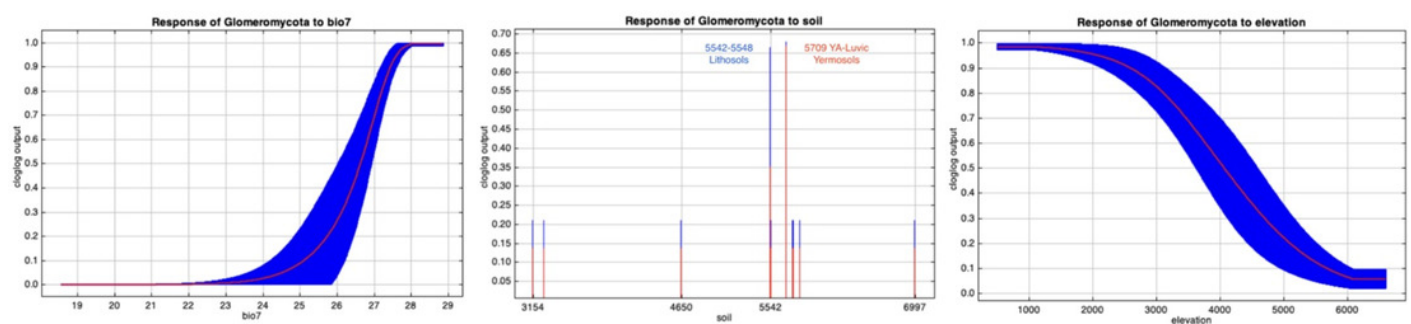
PeerJ

Table 1 (on next page)

Sum of weights of importance of all predictors as calculated from model weighting through Akaike Information Criterion and model-averaged coefficients.

A. Each value corresponds to the probability that a variable is included in the best predicted model. **B**. Conditional averaging resulting from model averaging using the function *dredge* in the R package MuMIn. Each estimate is tested for statistical significance as different from zero. (*: statistically significant).

1

A	Salinity	Organic matter	Elevation	pH	Assimilable P
Sum of weights	0.91	0.77	0.55	0.42	0.27

B	Estimate	Std.error	Adjusted SE	Z value	Pr(>|z|)	Signific.
(intercept)	-8.33502	4.64333	4.64382	1.795	0.0727	.
Organic matter	1.81588	0.70544	0.70587	2.573	0.0101	*
Salinity	-27.34102	12.52910	12.53072	2.182	0.0291	*
Elevation	1.19262	0.73985	0.74026	1.611	0.1072	
pH	-1.16042	1.07924	1.07971	1.075	0.2825	
Assimilable P	0.01587	0.66192	0.66237	0.024	0.9809	
Signif. codes: 0 *** , 0.001 ** , 0.01 * , 0.05 ·						

2

3

4