land

IMPACT FACTOR 3.2

CITESCORE 4.9

Article

Geodiversity as a Driver of Soil Microbial Community Diversity and Adaptation in a Mediterranean Landscape

Samuel Pelacani, Maria Teresa Ceccherini, Francesco Barbadori, Sandro Moretti and Simone Tommasini

MDPI

https://doi.org/10.3390/land14030583

land

MDPI

Article

Geodiversity as a Driver of Soil Microbial Community Diversity and Adaptation in a Mediterranean Landscape

Samuel Pelacani [1,*], Maria Teresa Ceccherini [2], Francesco Barbadori [1], Sandro Moretti [1] and Simone Tommasini [1]

[1] Department of Earth Science, University of Florence, Via G. La Pira 4, 50121 Florence, Italy; francesco.barbadori@unifi.it (F.B.); sandro.moretti@unifi.it (S.M.); simone.tommasini@unifi.it (S.T.)
[2] Department of Agriculture, Food, Environment and Forestry, University of Florence, Piazzale delle Cascine, 28, 50144 Florence, Italy; mariateresa.ceccherini@unifi.it
* Correspondence: samuel.pelacani@unifi.it

Abstract: Extreme meteorological events and anthropogenic influences determine important variations in microbial community composition. To know the extent of these variations, it is necessary to delve deeper into the geogenic factors to be considered as a baseline. The purpose of this study was to assess the effect of topographic characteristics and soil geochemistry on the spatial distribution of three Actinobacteria genera considered as molecular biomarkers of landforms belonging to Mediterranean environments. Given the important role that Actinobacteria play in the ecosystem, we performed a spatial distribution model of the genera *Rubrobacter*, *Gaiella*, and *Microlunatus* and investigated the fungi/bacteria ratio in a machine learning (ML)-based framework. Variable importance provided insight into the controlling factor of geomicrobial spatial distribution. The spatial distribution of the predicted Actinobacteria genera generally follows topographic constraints, mostly altitude. *Rubrobacter* was related to the slope aspect and lithium; *Microlunatus* was related to the topographic wetness index (TWI) and normalized difference water index (NDWI), as well as the fungi/bacteria ratio; and *Gaiella* was related to flow path and metals. Our results provide new information on the adaptation of Actinobacteria in Mediterranean areas and show the potential of using ML frameworks for the spatial prediction of OTUs distribution.

Keywords: geomorphometry; geomicrobiology; lanthanides; NDVI; machine learning; soil bacterial diversity; Actinobacteria; olive tree

Academic Editor: Gabriele Broll

Received: 16 January 2025
Revised: 22 February 2025
Accepted: 7 March 2025
Published: 10 March 2025

Citation: Pelacani, S.; Ceccherini, M.T.; Barbadori, F.; Moretti, S.; Tommasini, S. Geodiversity as a Driver of Soil Microbial Community Diversity and Adaptation in a Mediterranean Landscape. *Land* **2025**, *14*, 583. https://doi.org/10.3390/land14030583

1. Introduction

The importance of geodiversity as a determinant of microbial community structures in soils is still not well understood. Because geogenic influences are likely to develop over extended periods of geologic time, their effects may be masked by short-term changes in vegetation, land use, fires, or anthropogenic disturbances. However, these changes affecting the ecosystem could be predominantly driven by geogenic factors. Mediterranean landscapes are particularly vulnerable to climate changes that affect soil carbon sequestration capabilities, nitrogen supply, and phosphorus availability through variations in the bacterial community structure.

Well-preserved and seasonally managed soils with conservation practices of centuries-old olive orchards are an excellent archive for the Holocene, but the microbial DNA information of these archives has been neglected so far. According to recent studies, the soil of olive groves in the Mediterranean area is mainly dominated by the phyla Actinobacteria, Proteobacteria, Acidobacteria, Chloroflexi, Verrucomicrobia, and Firmicutes [1,2]. These

bacteria form biofilms, or assemblages of microbes, performing specific roles to promote the survival of the entire bacterial community.

In particular, Actinobacteria, filamentous gram-positive bacteria, play an important role in the ecosystem since they are involved in the decomposition of organic matter, bioremediation, and bio-weathering, producing a wide number of bioactive compounds [3,4]. Actinobacteria showed adaptability to many types of environments, known as extremophiles, including the higher temperatures of volcanic environments [5]; extremely cold temperatures, such as the Antarctic and Arctic regions [6]; and desiccated areas of the Atacama Desert [7]. Moreover, Actinobacteria are described as highly resistant to drought and to poor nutrient conditions and can positively influence plant growth by producing secondary metabolites with biocidal properties [8].

The rise in temperatures in the last decades that has affected the Mediterranean region—which is vulnerable to the impacts of warming, and especially to more prolonged and intense heatwaves and droughts than have previously been recorded [9]—has negative effects on soil biological activity [10]. Because of this, for a timely and adequate adaptation to the impacts of climate change and to strengthen resilience, it is crucial to know the relationships existing at the catchment scale between physiographic settings, soil geochemistry, and microbial composition and functions. Previous studies have found difficulty in assessing the positive or negative effect of climate on soil microbial functions, but the results potentially suggested a future homogenization of microbial communities by selecting specific taxa [11].

Empirical evidence and studies suggest that the paleoclimate and legacies of climate change influence both the resistance and resilience of soil bacteria and fungi abundance to extreme drought events, but not those of their community composition, richness, and multifunctionality [12,13]. Fungal community composition was shown to be more resistant but less resilient than the drought-related bacterial community [12]. It has been shown that the growth and carbon assimilation rates of soil microorganisms are influenced more by evolutionary history than by climate [14]. On the other hand, depending on their genetics and the physiological state in which they find themselves, soil microorganisms respond in different ways to environmental stress. Their ability to adapt to changes depends on the degree of perturbation and the time required for gene transcription and mutations. To date, modeling the physiological response to climate change remains a challenge [15]. Studying functional traits of the soil microbiome, such as the abundance of fast-growing microbes—opportunist "r-strategists"—versus the slow-growth "K-strategists" in relation to land use and the pedo-geomorphological characteristics of the landscape can help to predict how the soil microbiome will respond to different climate change scenarios [16].

The objective of this research is to investigate the bacteriome, considering it as a paleo-ecological archive. This is in order to deduce whether the bacteriome can resolve the natural and anthropogenic disturbances that occurred over the last centuries [17]. We hypothesized that the soil of ancient olive groves was inhabited by different microbial communities, and within them, Actinobacteria were selected by the changing environmental and geoclimatic conditions. By focusing on the hydro-geomorphometric factors of the landscape and the geochemical features of soils, we explored the possibility of highlighting some Actinobacteria as biomarkers.

Considering the vast panorama of studies on Mediterranean olive groves [18,19] and soil microorganisms as environmental stress indicators [20,21], to our knowledge, this study is the first to use a comprehensive set of geoinformatics, remote sensing, and geochemical tools to assess whether and how soil geodiversity influences the soil Actinobacterial community. Thus, we assessed the following: (i) whether soil microbial communities were influenced by landform and lithology; (ii) how geochemical parameters defined these influ-

ences; and (iii) whether such parameters can be predicted by remote sensing of vegetation and water indices.

2. Materials and Methods

2.1. Soil Sampling Strategy and Geochemical Analyses

Soil samples were collected during spring 2021 from 16 olive groves belonging to 6 geochemical environments of the Tuscany region, Italy (Figure 1): limestone, siltstone, sandstone, shale, volcanic rocks, and conglomeratic deposits. For these olive groves, the alfa and beta diversity of the bacterial community were assessed. The spatial distribution models of biomarkers were performed at the watershed scale, specifically for the Greve Basin located in the Chianti area, Florence (Figure 1B).

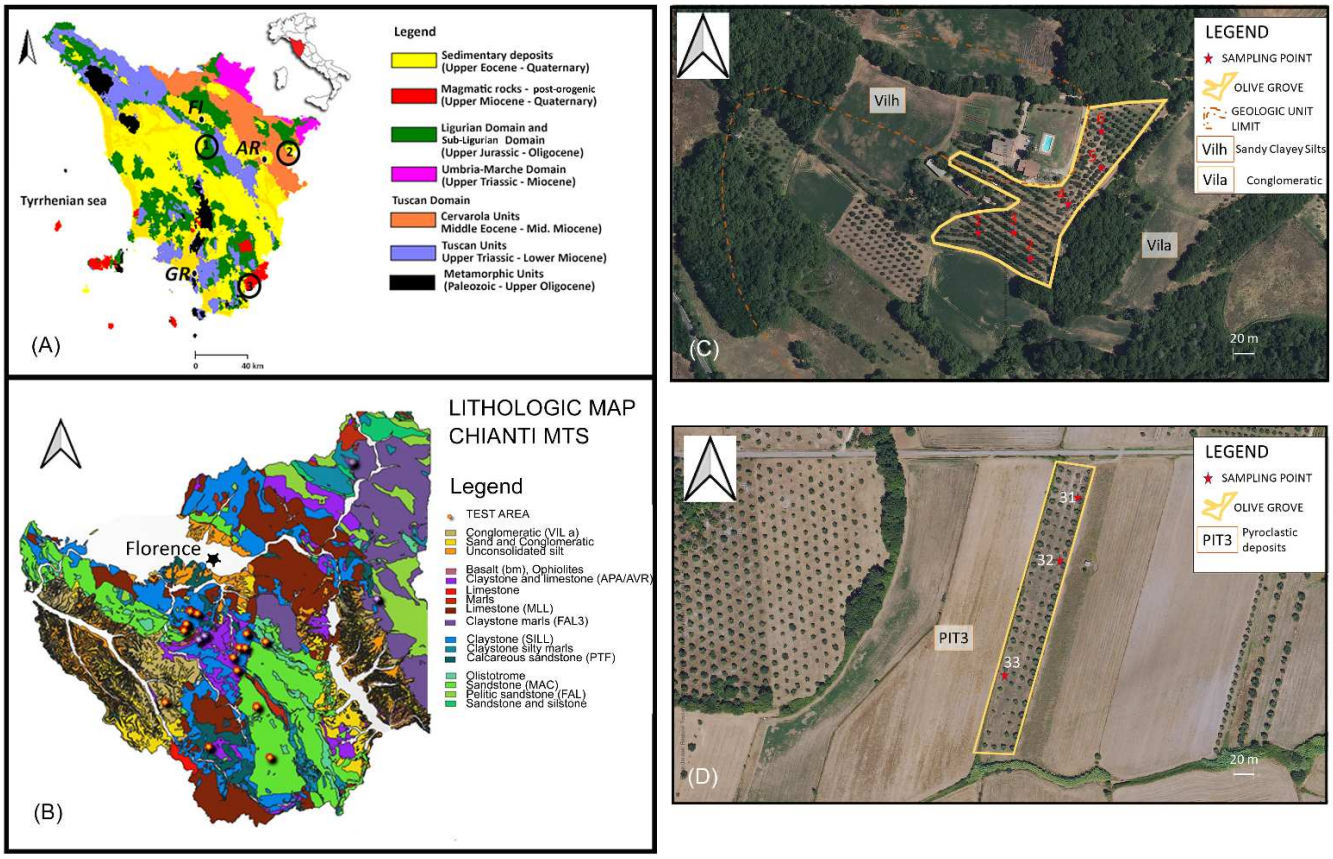

Figure 1. (**A**) Schematic geologic map of the Tuscany region, Italy. The different geochemical environments of the olive groves where the soil samples were collected are highlighted in black numbers, located in the following: (1) Chianti area (Florence—FI), (2) Alta Valtiberina (Aghiari, Arezzo—AR), and (3) Maremma (Pitigliano, Grosseto—GR); (**B**) Lithologic map of the Chianti Mts with the soil sampling sites (TEST AREA) collected in the Greve Basin; (**C**) Orthophotos (2023 year, Regione Toscana, Geoscopio WEB GIS) showing the sampling point distribution for the Anghiari olive grove; (**D**) Orthophotos (2023 year, Regione Toscana, Geoscopio) showing the sampling point distribution for the Pitigliano olive grove.

The study areas (Figure 1) were as follows: (i) Chianti, located in the central part of Tuscany, close to Florence; (ii) Alta Valtiberina, located in the eastern part of Tuscany, close to Anghiari (Arezzo); (iii) and Maremma, located in the southeastern part of Tuscany, close to Pitigliano (Grosseto). The olive groves are characterized by different landforms, lithology, and hence soil types (Table 1).

Table 1. Landform and sampling site descriptions for topsoil collected in 16 Tuscany olive orchards: soil texture; pH; SOM (soil organic matter); soil classification; lithology; aspect; and bioavailable contents of Mg, Li, Fe, Mn, Sb, and P for topsoil developed on 11 lithologies in Tuscany. The soil classification obtained from the Tuscany region soil database was referred to the World Reference Base (WRB—FAO, 2015). The subscripts beside the geological formation indicate the different sampling soils.

Landform	Site Location	Zone	Soil Samples	Soil Classification WRB	Lithology	Aspect	Sand (2 mm)	Silt (50 μm)	Clay (2 μm)	pH	SOM (%)	Mg [mg/kg]	Li [μg/kg]	Fe [μg/kg]	Mn [mg/kg]	Sb [μg/kg]	P [mg/kg]
Low gradient slope	Tosteto—Pitigliano	*Maremma*	31-32-33	Eutric Andosols	Pyroclastic deposits (Vulc—PIT)	SE	80.6	1.3	18.0	6.28	1.2	3.1	<0.1	101.0	2.3	<0.1	0.2
Graded	Faggeto—Anghiari (Arezzo)	*Alta Valtiberina*	1-2-3	Calcari Epileptic Cambisols	Fluvial-lacustrine deposits ($VILa_1$)	NE	46.1	42.6	11.3	7.41	1.5	34.8	14.1	22.3	1.7	0.37	12.9
	Faggeto—Anghiari	*Alta Valtiberina*	4-5-6	Calcari Epileptic Cambisols	Fluvial-lacustrine deposits ($VILa_2$)	NW	46.1	42.6	11.3	7.41	1.5	48.1	12.8	14.3	2.6	0.41	23.3
	Lamole—Greve in Chianti	*Chianti*	46-47-48	Eutric Cambisols	Macigno Formation (MAC)	S-SW	76.8	15.1	8.1	7.46	0.5	65.6	16.3	533.2	13.3	0.14	5.1
Midslope ridges	Torriano—Montefiridolfi, San Casciano	*Chianti*	13-14-15	Endoskeleti Calcaric Cambisols	Fluvial deposits (VILa)	NW	52.6	27.1	20.3	7.97	2.2	129.0	121.6	19.7	50.5	3.8	2.8
	Pruneti—Chiocchio,	*Chianti*	10-11-12	Calcari Endoleptic Cambisols	Palombini Shales (APA)	NE	26.4	42.4	31.2	8.15	1.6	208.0	97.4	17.0	3.0	0.87	0.3
	Pruneti—S. Polo in Chianti	*Chianti*	7-8-9	Calcari Endoleptic Cambisols	Marls (Marne)	SE	64.0	18.5	17.5	7.59	2.1	94.0	97.8	22.9	14.5	0.95	2.1
Upper slopes	Rignana—Greve in Chianti	*Chianti*	40-41-42	Eutri Epileptic Regosol	Pietraforte Formation (PTF)	S-SW	39.8	51.8	8.4	7.78	1.2	183.0	95.7	80.2	14.8	0.51	11.2
	Rignana—Greve in Chianti	*Chianti*	43-44-45	Calcaric Regosols	Varicolori Shales (AVR)	S-SW	29.5	66.0	4.5	7.62	1.3	33.0	45.8	27.7	9.1	0.6	2.2
	Castel Ruggero Monta Taurina	*Chianti*	37-38-39	Calcaric Regosols	Claystone (SIL)	S-SW	33.1	42.6	24.3	7.95	1.1	153.5	67.5	42.9	7.7	0.30	0.2
	Monteoriolo, Impruneta	*Chianti*	28-28-30	Calcaric Regosols	Claystone (SIL)	E	61.3	22.5	16.2	7.05	2.6	51.6	<0.1	385.5	4.3	1.5	7.5
	Pruneti—I Tinti, Strada in Chianti	*Chianti*	16-17-18	Calcari Endoleptic Cambisols	Basalts (bas)	N-NW	61.9	31.5	6.6	7.90	2.3	408.8	24.12	43.7	3.9	1.7	2.0

Table 1. *Cont.*

Landform	Site Location	Zone	Soil Samples	Soil Classification WRB	Lithology	Aspect	Sand (2 mm)	Silt (50 μm)	Clay (2 μm)	pH	SOM (%)	Mg [mg/kg]	Li [μg/kg]	Fe [μg/kg]	Mn [mg/kg]	Sb [μg/kg]	P [mg/kg]
Open slopes	Pruneti—Lizzano	*Chianti*	22-23-24	Calcaric Regosols	Monte Morello Formation (MLL)	N-NW	36.7	38.9	24.4	8.07	1.2	98.7	41.1	94.6	5.8	0.6	1.4
	La Querce—Impruneta	*Chianti*	25-26-27	Calcaric Regosols	Monte Morello Formation (MLL)	W	22.7	68.4	8.9	8.12	1.1	132.8	9.3	91.5	4.9	0.7	1.6
	Erta di Quintole—Impruneta	*Chianti*	19-20-21	Endoskeleti Calcaric Cambisols	Carbonatic flysch (Flysh)	S	57.6	22.4	20.0	7.54	2.8	235.8	76.5	128.7	7.6	1.1	3.3
	Castel Ruggero Poggio Fontaccia	*Chianti*	34-35-36	Calcaric Regosols	Varicolori Shales (AVR_1)	S-SW	34.5	51.3	14.2	8.09	1.1	119.2	79.1	104.7	3.1	0.36	1.7

Only traditional and organic olive grove management cultivations are taken into account. In terms of farming management, the most common regime is the traditional rainfed olive grove, characterized by about 300 trees per hectares. The median surficial areal extension of single olive groves is about 0.9 hectares. For each olive grove, three samples of the topsoil-associated rhizosphere (<30 cm deep) were collected over a period of one working week. For each olive plant, three soil samples were collected and homogenized to obtain a representative one. The soil samples were transported to the laboratory in cool boxes. Keeping the humidity around 4%, each replicate of soil was gently sieved at 2 mm and kept at −20 °C until DNA extraction. For metagenomic analysis, the samples were stored at −20 °C until DNA extraction. The selected olive groves lie on geomorphodynamically stable landforms where the soil developed from its parent rock material and is not subject to soil erosion or depositional processes of sediments. Therefore, samples were collected from olive groves lying on the upper part of the relief, where the soil horizons develop from each selected bedrock. If these conditions were not satisfied, the soil hillslope catena concept was taken into consideration for sampling design.

The bioavailable content in the <2 mm fraction of each soil sample of major elements (Al, Ca, Fe, K, Mg, Na, and P), heavy metals (Cd, Pb, Cu, Ni, As, Cr, Ba, Co, Mn, Se, and V), lanthanides (REEs plus Y and Sc), and other trace elements (Li, Sb, Te, Be, Tl, Sn, Ti, Pd, and Mo) were analyzed by quadrupole inductively coupled plasma mass spectrometry (ICP-MS; Agilent Technologies 7900 ICPMS—Hachioji, Tokyo, Japan) equipped to remove polyatomic interference. The reproducibility of measurements was assessed by combining an internal standard normalization by rhodium and an external calibration using the AGV-1-certified reference materials. The accuracy (RSD) for all of the estimated elements varied from 2 to 12.9% at the μg g^{-1} levels of concentration. The variability of the analytical repetitions was <10% at the ng kg^{-1} levels of concentration. The soils were treated with DTPA at pH 5 to simulate the uptake by the rhizosphere microbiome (Ehlers and Luthy, 2003; Feng and Shan, 2005). Each sample was determined in triplicate. The bioavailable fraction was extracted by reacting 10 g of dried soil with 20 mL 5 mM DTPA solution at pH 5. The obtained suspension was stirred for 24 h at 25 °C, filtered with Millipore™ membranes (0.45-μm membrane filter), and diluted 50 times.

REEs include the 15 lanthanides (atomic numbers 57–71), and according to their physical and chemical properties, they are usually divided into light (L-REEs = light rare earth elements from La to Sm), medium (M-REEs from Sm to Gd), and heavy (H-REEs from Gd to Lu) REEs.

2.2. Geological Settings

The geology of Tuscany encompasses numerous tecto-sedimentary facies belts and younger basins bordered by metamorphic core complexes associated with thirteen paleogeographic domain and geotectonic units, outcropping of the nappe stack of the Northern Apennines (Figure 1). These tectonic units derived from the Apulia microplate, continental margin, and the Tethys seafloor sediments were emplaced during Tertiary subduction and continental collision.

From the bottom, the following are exposed: the Tuscan metamorphic units, the Tuscan nappe, the Sub-Ligurian units, the Ligurian units, the epi-Ligurian succession, and the post-orogenic sedimentary succession starting from the Middle Miocene to the Quaternary.

In this study, we focused on the following three important territories for olive grove production in the Tuscany region (Figure 1A): (i) Chianti, located in the central part of Tuscany, close to Florence, and characterized by five non-metamorphic tectono-sedimentary units and Plio-Pleistocene fluvial deposits; (ii) Alta Valtiberina, located in the eastern part of Tuscany, close to Anghiari (Arezzo), and characterized by Pliocene-Quaternary fluvio-

lacustrine deposits; (iii) and Maremma, located in the southeastern part of Tuscany, close to Pitigliano (Grosseto), and characterized by magmatic rock related to the opening and spreading of the Northern Tyrrhenian Sea back-arc basin during the late Miocene to the Pleistocene period.

In the Chianti area (Figure 1B), the widest outcroppings characterized by olive groves are the sediments belonging to the external Ligurian domain that were thrust onto the upper Oligocene–lower Miocene Macigno sandstones (Tuscan units) deposited in the foredeep basin. The Macigno sandstone formations (MAC) represent the uppermost portion of the Tuscan units, with a thickness of up to 1500–2000 m of siliciclastic turbidite sandstones and siltstones, conglomerates, and claystone. The sandstones are graywackes dominated by metamorphic rock fragments followed by volcanic and sedimentary rocks. Furthermore, the Macigno Formation hosts interbedding up to 50 m thick of lenticular successions of hemipelagic marls and very fine/distal turbidites (San Polo Marls—"Marne").

The Ligurian units consist mainly of the following four formations, starting from the bottom: (i) Palombini shales (APA; Cretaceous)—highly tectonized gray claystones with interbedded fragments of ophiolitic complexes (basalts—bm, gabbros, serpentinite breccias, and reddish cherts) representing the terrigenous cover of the NW Tethyan ocean crust; (ii) Pietraforte formation (PTF; upper Cretaceous)—turbiditic sandstones alternated with claystones, including calcareous and dolomitic rock, plagioclase, and quartz. This formation is related to a turbiditic system accumulated in the oceanic trench during the subduction of the NW Tethys; (iii) Sillano formation (SILL; upper Cretaceous–Paleogene)—a chaotic body of mixed rocks including blocks of different ages and origin such as the varicolored shales (AVR, green, gray, brown, and red) alternating quartz-rich or carbonaceous turbiditic sandstone and siltstones, marls and calcareous marlstone, plagioclase, and phyllosilicate; and (iv) Mt. Morello formation (MLL; Early to Middle Eocene)—micritic limestones, bioclastic calcarenites, marlstones, and claystones, with arenites containing carbonate cement and quartz, weathered plagioclase, and mica. A post-orogenic fluvial deposit was also considered in the Chianti area (VILa; Plio-Pleistocenic deposits).

The lithology belonging to the Middle Pleistocene monogenic conglomerates (VILa) corresponding to the top of the fluvio-lacustrine deposits of the Anghiari-Sansepolcro Basin was selected in Alta Valtiberina, Arezzo (Figure 1C), and it is part of the larger Tiberino Basin.

The Pitigliano Fm. (Maremma, Figure 1D) represents the youngest deposit of the Latera volcanic sequence, and it is characterized by a complex sequence comprising airfall, pumice deposit, flow deposit, strongly welded ignimbrite, and pyroclastic flow containing a glassy matrix.

2.3. DNA Extraction, qPCR Quantification, and Sequencing of 16S rRNA Gene

The microbial composition was determined by extracting the total genomic DNA from 500 mg of soil by FastDNA™ SPIN Kit for Soil (MP Biomedicals, Santa Ana, CA, USA) [22]. The quality control and DNA yield were checked by 1% agarose gel electrophoresis and by spectrophotometer (Picodrop limited, Hinxton, UK).

A quantitative PCR test was performed to determine the 16S rRNA gene copy number of bacteria and the 18S rRNA gene copy number of fungi $\times$ g^{-1} of soil using 40 ng DNA templates for all the samples. Reactions were performed in an iCycler (Bio-Rad, Herculesc, CA, USA), and the results were analyzed with the manufacturer's software (Optical System Software v 3.0a). Amplification was carried out in a 25 µL final volume containing the following: 2.5 pmol of each primer, 12.5 µL of iQ SYBR Green Supermix (2X), and sterile ddH_2O to reach the appropriate volume; three replicates were carried out for each sample. Amplification reactions were performed in 96-well microtiter plates (BioRad) with a known amount of

Bacillus subtilis BD1512 341f/515r 174 bp PCR fragment and of *Saccharomyces boulardii* (Zambon Italia, Bresso, Italy) FF390/FR1 390 bp PCR fragment (both previously amplified and purified) in each plate, used to develop the standard curve for the respective qPCRs by plotting the logarithm of known concentrations (from 10^{-1} to 10^{-6} ng in 25 µL) against the threshold cycle (Ct) values. The qPCR program for eubacteria had an initial step of denaturation (3 min, 95 °C) followed by 40 cycles of 15 s at 95 °C, 30 s at 63 °C, and 30 s at 72 °C; for fungi, an initial step of denaturation was performed (3 min, 95 °C) followed by 40 cycles of 45 s at 95 °C, 30 s at 50 °C, 50 s at 70 °C, 25 s at 90°C, and 4 min at 72 °C. After each cycle, a melting curve program was run, for which measurements were made at 0.5 °C temperature increments every 10 s within a range of 60–100 °C [23].

NGS sequencing and library preparation were performed by Novogene Company (Novogene, Beijing, China) on the 16S rDNA V3-V4 region on NovaSeq PE 250 System.

2.4. Bioinformatic Analyses

Paired-end reads were assigned to samples based on their unique barcodes, truncated by cutting off the barcode and primer sequences, which were subsequently merged using FLASH (V1.2.7) [24]. Quality filtering on the raw tags was performed according to Qiime process (V1.7.0) [25,26]. The tags were compared with the reference Gold database using the UCHIME algorithm [27] to detect and remove the chimera sequences [28], obtaining the effective tags.

Then, microorganism sequences were classified with a threshold at ≥97% using the Uparse software (Uparse, v7.0.1001) [29]. For each representative sequence, Mothur software was performed against the SSU rRNA database of the SILVA Database [30] for taxonomic species annotation.

To obtain the phylogenetic relationship of all microorganism representative sequences, the MUSCLE alignment software was used (Version 3.8.31).

In order to identify differences in the bacterial community structures among the different geochemical environments, the β diversity was estimated based on Bray–Curtis dissimilarities between topsoil samples, and then non-metric multidimensional scaling analyses (NMDS) was performed to explain the pairwise dissimilarity between objects in a low-dimensional space.

Bacterial species richness within a single geochemical environment, that is, the α diversity indices of the soil bacterial community, was calculated by the Chao1 index, which is sensitive to rare species [31], and the Shannon index. Comparison of the abundance differences of species between groups (topsoil developed in different geochemical environments) under different taxonomic levels was calculated by MetaStat tests. Diversity in the bacterial community related to different geochemical environments was evaluated considering the bioavailable fraction of REEs and performed by principal component analyses (PCA). The PCA was applied in covariance modes that indicate the direction of the linear relationship between variables and how they vary and in correlation mode that measures both the strength and direction of the linear relationship between variables (OriginPro, Version 2024. OriginLab Corporation, Northampton, MA, USA). In this study, the REEs were used as a tool to trace the natural geochemical evolution processes of different geological settings, mostly sedimentary systems, that differ for their biogeochemical and microbial signature. The present study used information on the spatial distribution of REEs and their relationship with different geochemical environments and bacterial composition generated by the authors' previous studies [2,17].

2.5. Geoinformatic Analyses

The basic idea was to find automatically selected terrain covariates using a machine learning (ML) framework to analyze the relationships between the hydro-geomorphological feature of the landscape and the bacterial composition and diversity of the study area. We hypothesized that microbial diversity can be traced, and hence predicted, using REEs geochemical composition, which in turn is strictly related to landforms defined as a distinctive feature of the land surface shaped by erosion, accumulation, or deformational processes. More in depth, the prediction modeling was focused on the following main genera of the Actinobacteria spatial distribution found in the topsoil of the Chianti area (Florence, Italy): *Rubrobacter*, *Gaiella*, and *Microlunatus*. The Chianti area was selected for this type of investigation because of its higher number of sample points available at the sub-basin scale compared to Alta Valtiberina (Anghiari) and Pitigliano (Grosseto), other areas studied in this research.

Geomorphometric covariates, widely used in the soil spatial modeling framework, were selected as drivers of Actinobacteria spatial distribution and bacterial community characterization. Using SAGA GIS, we applied a detailed digital terrain analysis to quantify and describe the terrain landforms through topographic indices extrapolated from a 12 m × 12 m cell-size-resolution DTM [17] obtained from the TanDEM-X radar imaging satellite platform provided by the Deutsches Zentrum für Luft und Raumfahrt (DLR). The DTM was preprocessed in order to extract artefacts and to guarantee hydrological functionality (fill sinks) using the algorithm proposed by Wang and Liu [32]. Subsequently, the following terrain indices were derived from the preprocessed DTM: (i) three basic terrain derivatives (slope, aspect, and catchment area); (ii) topographic wetness index (TWI) based on modified catchment area and slope tangent [33]; (iii) valley depth (VD); (iv) channel network base level (CNBL); (v) vertical distance to channel network (CND) [34]; (vi) length slope factor (LS factor); (vii) topographic position index (TPI) [35], defined as the difference between a central pixel and the mean of its surrounding pixels; and (viii) terrain ruggedness index (TRI), the mean of the difference between the central pixel and its surrounding pixels [36]. The normalized difference vegetation index (NDVI) and the normalized difference water index (NDWI) were derived from remote-sensing data. The NDVI is useful to understand the distribution of vegetation and provides information about the relation between vegetation cover and soil. The NDWI is also a remote sensing-based indicator sensitive to changes in the soil water content or in the water content of leaves. The NDVI and NDWI data were extrapolated from Advanced Spaceborne Thermal Emission and Reflection Radiometer (Sentinel-2) images. One image from 27 June 2020 was downloaded from USGS Earth Explorer and corresponds to the same period of the soil sampling survey. The NDVI was calculated applying the following formula: NDVI = (Band 8 − Band 4)/(Band 8 + Band 4). The NDWI was calculated applying the following formula: NDWI = (Band 3 − Band 8)/(Band 3 + Band 8).

The effect of lithology on the spatial distribution at the basin scale of Actinobacteria main genera was determined by using two fractionation ratios of soil bioavailable REEs and olive fruit REEs, La/Sm and La/Yb, that represent Light-REE/Medium-REE and L-REE/Heavy-REE. These fractionation ratios are a powerful ecological index to evaluate and monitor the system dynamic and the related evolutionary trends at the basin level, which has been evaluated in our previous study [17].

2.6. Machine Learning and Geostatistical Modeling

We used a combined approach including multiple geostatistical models to address our hypothesis. In particular, we used (1) random forest (RF) analysis to discover key predictors of bacterial community richness and composition, including geomorphometric

spatial predictors, soil geochemistry, and remote sensing indices such as the vegetation index (normalized difference vegetation index, NDVI) and the water index (normalized difference water index, NDWI). Hence, RF was used to identify whether the geogenic factors can explain a unique portion of the variation composition of the bacterial community that cannot be represented by other key predictors of soil microbial communities; (2) multiple regression analyses (MLRA) [34,37] were used to perform a spatial prediction of a bacteriome distribution model belonging to the main genera of Actinobacteria for a Mediterranean landscape. All of these geoinformatics and statistical models address a particular part of our research to answer our question that cannot be addressed using each approach individually.

The ML makes it possible to investigate the non-linear relationship between the dependent variable and independent ones. The RF model is an ensemble model that uses a collection of decision trees to predict the target value and can be used for both classification and regression models. The RF works by randomly selecting subsets of input features (feature sampling) and subsets of training data (bootstrap sampling) for building each decision tree. The final prediction is made by aggregating the output of all trees, calculating the mean for regression tasks. During the training phase, RF also calculates the out-of-bag-error (OOBE), which estimates the model's performance using data not included in the bootstrap sampling and providing unbiased estimation of the model's accuracy. Finally, the model's performance was evaluated using the root mean square error (RMSE). The correlation between predicted and measured values of dependent variables was retrieved using the R2 value. The optimal initialization parameters for the RF model to obtain a satisfactory RMSE are summarized in Supplementary Table S1. Moreover, the sensitivity of the chosen parameters was assessed.

Individual MLRA models were performed for OTU2/OTU3 ratio (*Rubrobacter*/*Microlunatus*), OTU354/OTU2 ratio (*Gaiella*/*Rubrobacter*), OTU354/OTU3 ratio (*Gaiella*/*Microlunatus*), and fungi/bacteria (F:B) ratio based on local geomorphometric and geochemistry variables. These distribution patterns support the hypothesis that the ratios represent the complex interactions of all the OTUs with each other (inside the bacterial community) and with the different geomorphological and geochemical settings. In this way *Rubrobacter*, *Microlunatus*, *Gaiella*, and the F:B ratio were considered as bioindicators.

The interpolation of the data points and their spatial regression were performed in SAGA GIS [34] using a geoprocessing tool: multiple regression analyses. The independent and dependent variables were converted into a raster format. The raster grid centroids were transformed into points, and at each point, the respective values of the independent variables were added to the attribute table in SAGA GIS, resulting in a dataset of 1,800,821 points for predicting the spatial variation of bioindicators at the Greve Basin level.

Independent variables (N = 16) were selected using the forward stepwise selection procedure, and a regression model applied to predictor grids with interpolated residuals added was chosen. The models were built using as resampling method B-spline interpolation, including X and Y coordinates, as a significance level of $p < 0.05$ and a 10-fold cross validation subsample was used. Model validation was first performed by comparing observed and model-predicted values using the coefficient of determination (R2) and R bar squared of the multiple linear model, R2 adj, as reported in Supplementary Table S2. Then, the R2 adj of the final model was analyzed after the residual's calibration by regression. Finally, the quality of the maps was assessed by comparing observed values with estimated values at each sampling point. The R2 adj allows the comparison of regression models that contain different numbers of predictors. Similar to the R-square, it measures variance explained by the model, but it adjusts using the residual degree of freedom. It is generally the

best indicator of the fit quality to compare statistical models. R2 adj is always ≤R-square, with a value closer to 1 indicating a better fit.

The MLRA models help to understand the following: (i) which environmental variables affect the distribution of the selected bacteriome in the soils of the olive groves; (ii) how topographical, geochemical, and hydrological variables are linked (geodiversity); and (iii) how bacteriome ratios vary in soils according to geodiversity. The performance of the MLRA models is reported in the Supplementary Material (Supplementary Table S2).

The *Actinobacteria*-representative OTUs (*Rubrobacter*, *Microlunatus*, and *Gaiella*) for the olive grove soils under consideration were discovered by applying a non-parametric decision tree learning technique (CART, Salford Predictive Modeler V. 8.3.4). These target OTUs were predicted from a database of 9864 OTUs [2] and considering 40 metals, including the lanthanides as covariate, where the lithology classes were selected as initial target variables.

3. Results

3.1. Geomicrobial Diversity Characterization and Community Composition for Ancient Olive Groves of the Tuscany Region

We initially examined the relationship between soil geochemical characteristics and the diversity of bacterial community structure using the soil bacterial alfa and beta diversity indexes. Our results indicated different diversity patterns (Shannon index) based on the geochemical environmental: higher for pyroclastic-derived soil and lower for conglomeratic-derived soil (Figure 2A). A significant difference in the bacterial richness (Chao1) between the claystone lithologies and limestone was observed, which revealed a notable higher bacterial richness in varicolored claystone compared to the Palombini shales or conglomeratic soils (Figure 2B).

Further, the beta diversity projection on the NMDS plot revealed that the pyroclastic and arenaceous soils' bacterial communities were significantly different (Figure 2C) among themselves and with respect to other soils. To further confirm the difference of bacterial community composition at the phylum level, the MetaStat method was used to screen the bacteria species (OTUs) with significant differences, and a heat map was drawn (Figure 2D). The results showed that there were significant differences in the bacterial communities between pyroclastic derived soil and the other topsoil. The relative abundance of Actinobacteria in the pyroclastic-derived soils was significantly lower than in the conglomeratic-, basaltic-, and carbonatic flysh-derived soil of the Chianti area (Group E, F, G in Figure 2D; Supplementary Figure S1); conversely, in pyroclastic soil, levels of *Gemmatimonadetes* and *Acidobacteria* were significantly higher than in soils developed on sandstone, conglomeratic, basaltic, and carbonatic flysh lithologies (Figure 2D; Supplementary Figure S1).

Using REEs as a proxy for soil geochemical characteristics, the results indicated that higher REE content was able to select for a soil bacterial composition that differed from those soils with the following: (i) high sand fraction; (ii) high clay fraction; and (iii) skeletal soil (Figure 3A).

A first exploratory analysis using regression tree machine learning techniques highlighted the presence of four bacteriomes (OTUs) among the main genera of Actinobacteria discretizing the different studied geochemical environments: OTU95 (Actinobacteria, unknown), OTU2 (*Rubrobacter*), OTU3 (*Microlunatus*), and OTU354 (*Gaiella*). This information was used to discover the existence of a linear relationship between significative fractionation ratios of REEs and bacteriome (Figure 3B; Table S3). Four geomicrobial clusters were identified using PCA that identified associations and gradients of Actinobacteria and REEs for the claystone, varicolored shale, conglomeratic deposits, and limestone environ-

ments (Figure 3B). *Microlunatus* in association with OTU95 showed linear correlation with pyroclastic soils (point samples 31, 32, and 33; Figure 3B).

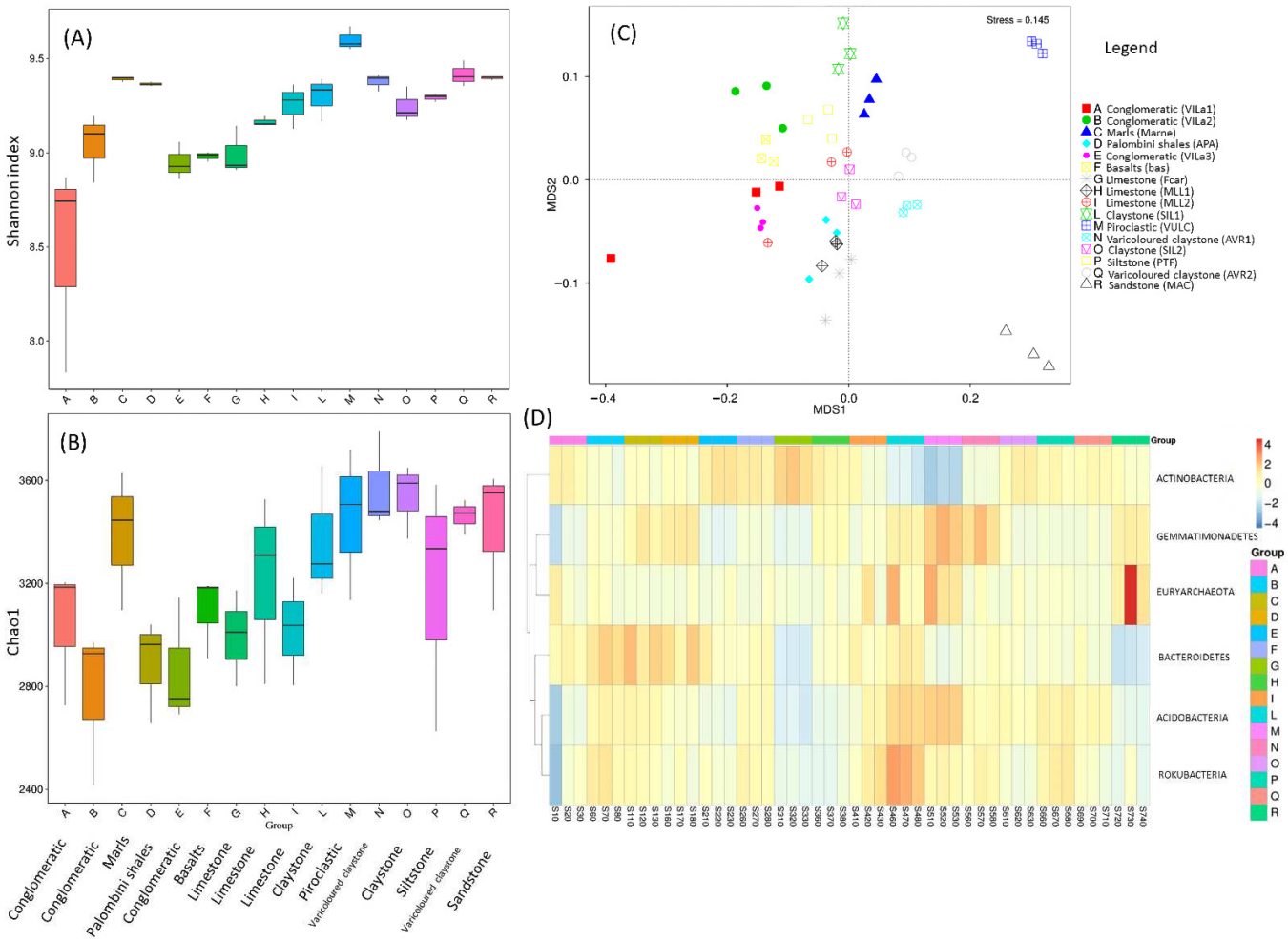

Figure 2. Biodiversity of the bacterial community in contrasting Mediterranean geochemical environments. Differences were observed in the bacterial Shannon index (**A**) and OTU richness (Chao 1, (**B**)) among different olive groves. The top and bottom boundaries of each box indicate the 75th and 25th quartile values, respectively, and lines within the boxes represent the median values. (**C**) Nonmetric multidimensional scaling (NMDS) of bacterial communities among topsoil based on Bray–Curties dissimilarities. (**D**) Heat map of bacterial communities at phylum level of olive grove soils under different Mediterranean geochemical environments.

3.2. Relative Importance of Hydro-Geomorphometric and Geochemistry Factors on Soil Actinobacterial Diversity

To test whether local and landscape environmental heterogeneity characteristics affect Actinobacteria distribution in soils, we applied a random forest (RF) model. We used the RF model to assess the variable importance and, hence, to reduce complexity, selecting the most important variables. The results of the RF analysis revealed that geomorphometric factors accounted for a different proportion of the variance in soil Actinobacteria (Figure 4). For *Microlunatus* (OTU3), the influence of their interactions with the other Actinobacteria and geomorphometric descriptors was stronger than the influence of geochemical spatial descriptors (Figure 4). Conversely, for *Gaiella* (OTU354), microbial communities were associated with geochemical covariates such as the Sb (antimony) chalcophile element above 45%, the Mo (molybdenum) siderophile element higher than 35%, and the Al (aluminum) lithophile element above 30%. Considering soil nutrients such as phosphorus was less

influential, with 13% of the explained variance. The geomorphometric predictors CND (84%) and elevation (DEM 49%) showed higher predictive importance for *Gaiella* (OTU354) than *Rubrobacter* and *Microlunatus* genera. The hydro-morphological characteristics of the soil landscape, TWI (49%), TPI (30%), and NDWI (11%), showed higher importance for *Microlunatus* (OTU3). Moreover, this genus was highly related to the multispectral reflectance of the vegetation (NDVI > 40%). RF also indicated a strong link between the *Rubrobacter* genera (OTU2), *Microlunatus*, and *Gaiella*, suggesting a co-occurrence network. In fact, OTU2 showed a higher importance variable value (from 64% to 82%) for predicting *Microlunatus* and *Gaiella*, respectively. Moreover, for OTU2, the lithophile elements, Li (Lithium) and Ni (Nickel), showed an importance of 28% and 17%, respectively. Furthermore, the occurrence of the OTU2 is connected to the presence of the chalcophile elements strontium (17%) and calcium (15%).

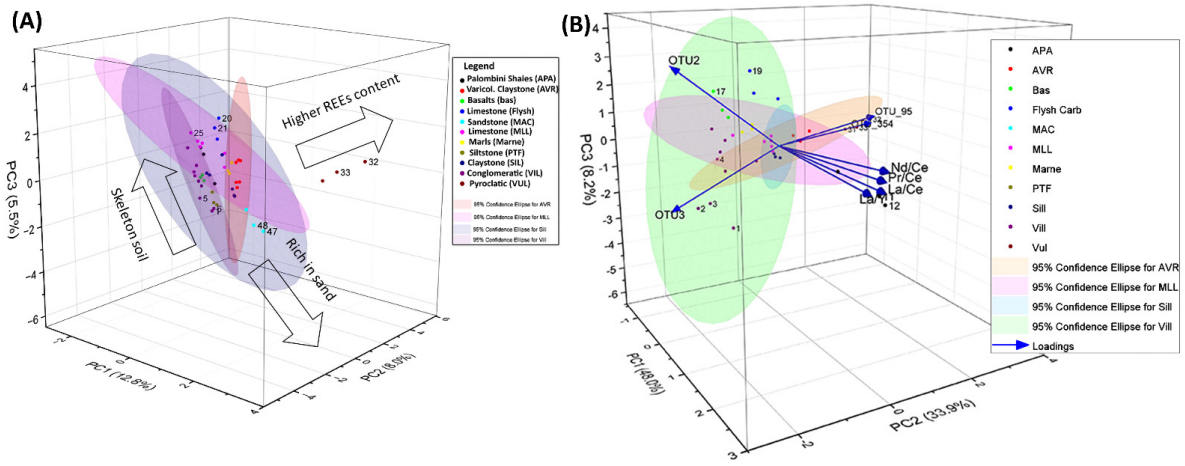

Figure 3. (**A**) PCA analysis for the 48 topsoil microbial communities and bioavailable fractions of REEs in ancient olive groves cultivated in the Tuscan landscape grouped by geology units (**B**). PCA analysis for 48 topsoil Actinobacteria (OUT_95; *Rubrobacteria* OTU2; *Microlunatus* OTU3; *Gaiella* OTU354) and representative fractionation parameters of light, medium, and heavy REEs (La/Ce), (Pr/Ce), (Nd/Ce), (La/Y).

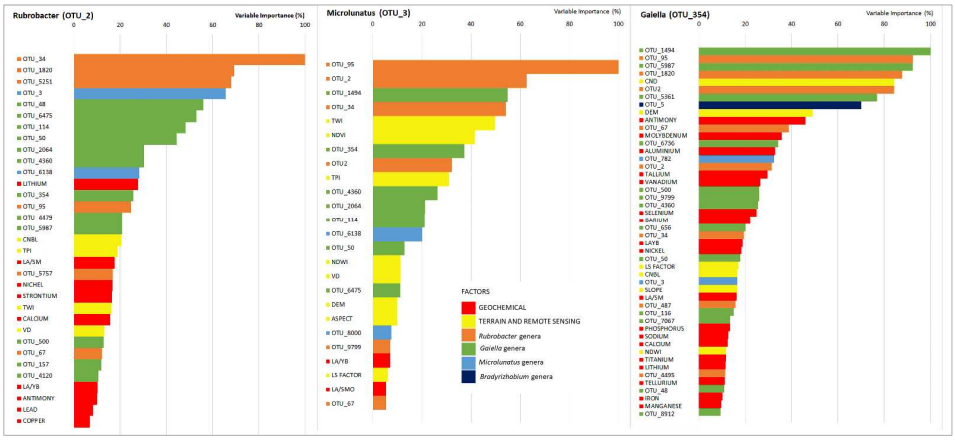

Figure 4. Random forest-predicted importance of the studied variables (geomorphometric; geochemical factors; remote sensing; and *Rubrobacter*, *Microlunatus*, and *Gaiella*) for soil Actinobacteria OTU2, OTU3, and OTU354, characterizing the landform of the Greve Basin (Florence, Italy). Legend: channel network base level (CNBL), valley depth (VD), vertical distance to channel network (CND), terrain ruggedness index (TRI), topographic position index (TPI), topographic wetness index (TWI), digital elevation model (DEM), length slope factor (LS factor), normalized difference vegetation index (NDVI), and normalized difference water index (NDWI).

3.3. Spatial Model Prediction of Biomarkers Based on the Most Dominant Actinobacteria Genera Across an Ancient Olive Grove Mediterranean Landscape

To investigate the co-occurrence network of the selected OTUs in relation of the environmental variables, we modeled the spatial distribution of the OTU2/OTU3 ratio (*Rubrobacter*/*Microlunatus*), the OTU354/OTU2 ratio (*Gaiella*/*Rubrobacter*), the OTU354/OTU3 ratio (*Gaiella*/*Microlunatus*), and the fungi/bacteria ratio (F:B), considered a biomarker for the Greve Basin (Figure 5).

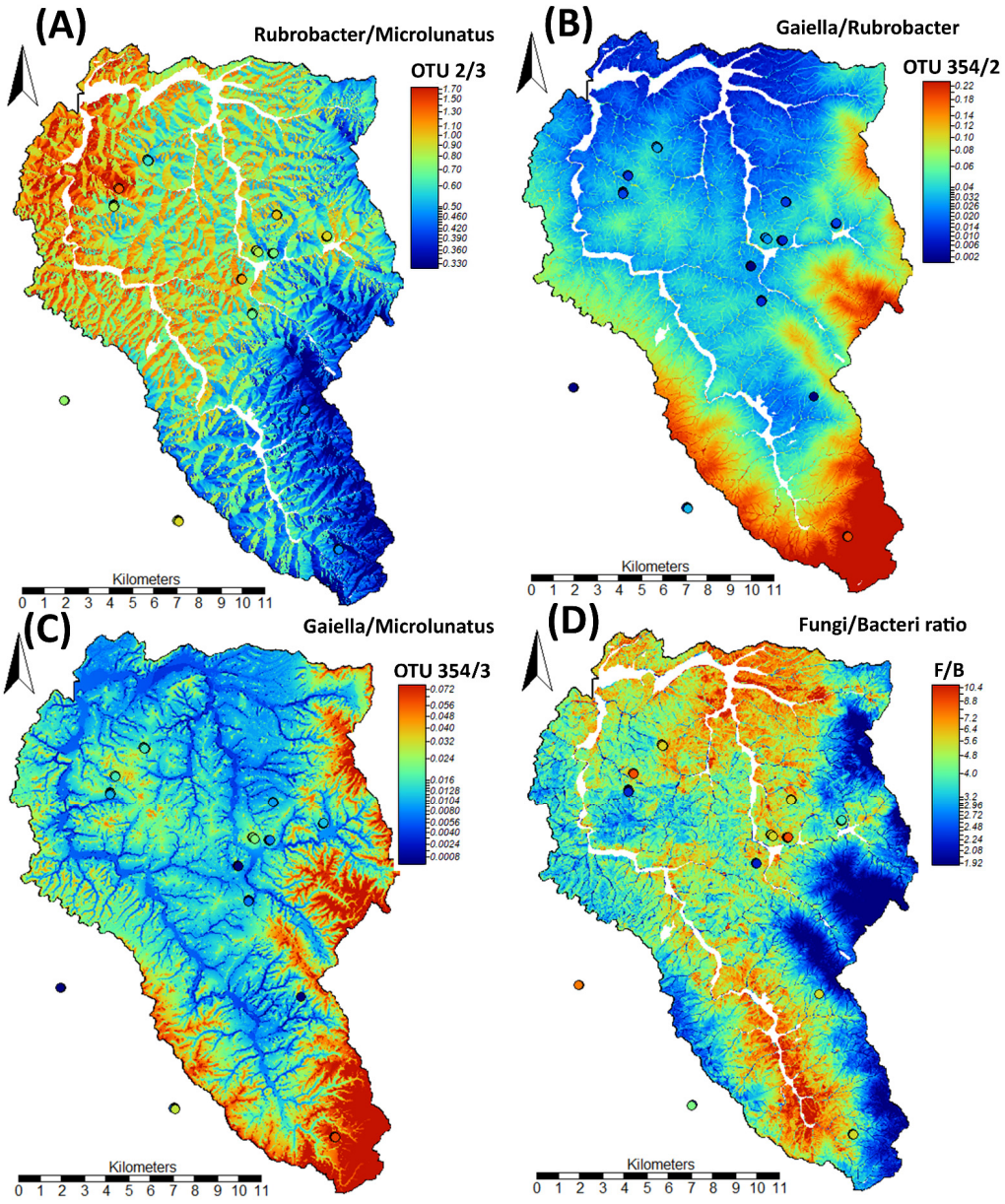

Figure 5. Spatial prediction of biomarkers: (**A**) *Rubrobacter*/*Microlunatus*, (**B**) *Gaiella*/*Rubrobacter*, (**C**) *Gaiella*/*Microlunatus*, and (**D**) fungi/bacteria ratio for a Mediterranean landscape dominated by ancient olive groves, Greve Basin, Tuscany, Italy.

The higher OTU2/OTU3 ratio (*Rubrobacter*/*Microlunatus*) values were found in the northwestern part of the basin, characterized by limestone soils, rather than the sandstone outcropping in the southern part (Figure 5A). The ratio values of the biomarkers OTU2/3 (*Rubrobacter*/*Microlunatus*) ranged from 0.32 to 1.72 in sandstone and calcium carbonate limestone-derived soil, respectively. The model results revealed that high values of OTU2/OTU3 ratio were mainly related to south-facing slopes. In agreement with the

MLRA, the model statistics revealed that the aspect explained 24.4% of the variance in the dependent variable OTU2/3 (Figure 6). MLRA analyses showed that aspect, longitude (X), LS factor, and elevation were the significant predictors of spatial variation in bacterial OTU2/3 (Figure 6A) for the Greve Basin.

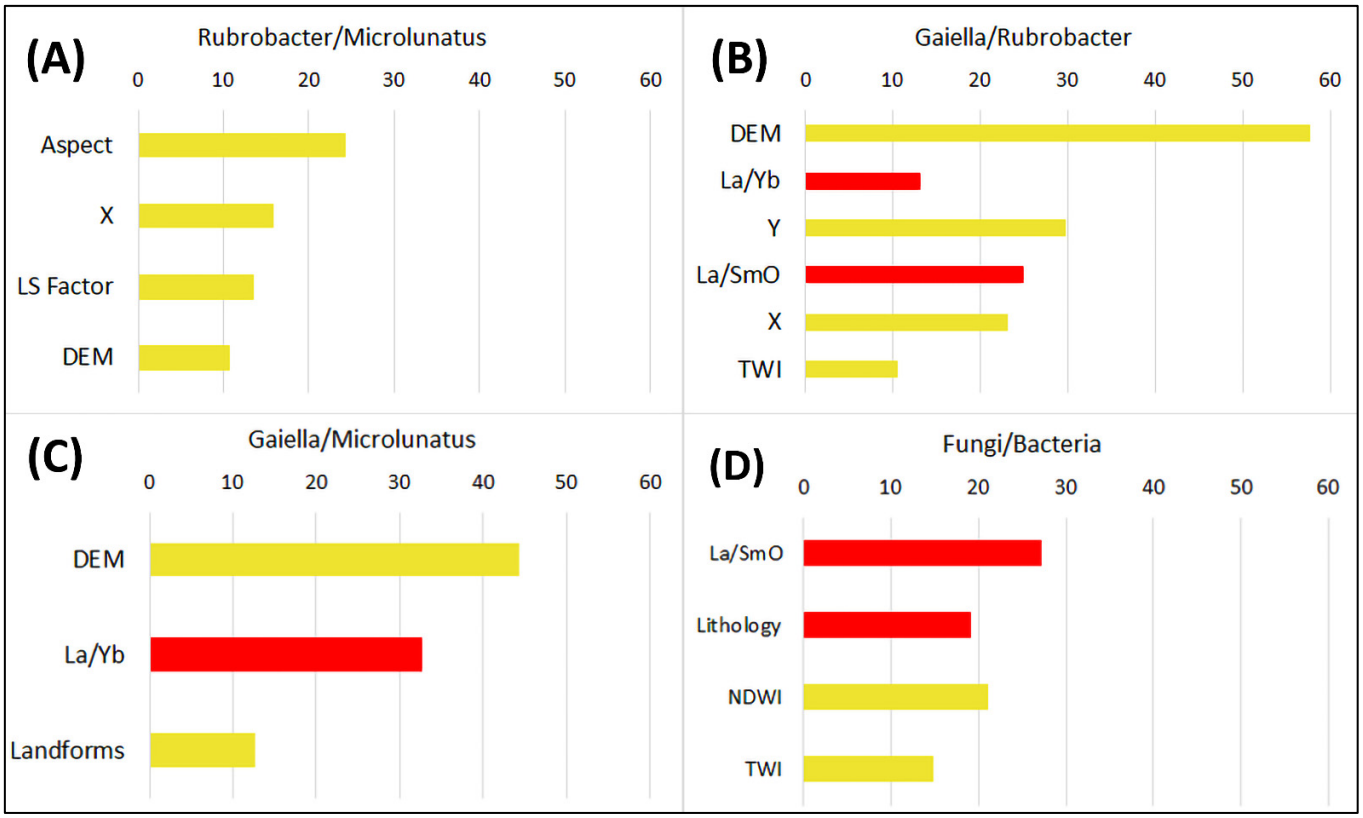

Figure 6. Multiple regression analysis predicted importance of soil Actinobacteria: (**A**) *Rubrobacter/Microlunatus*, (**B**) *Gaiella/Rubrobacter*, (**C**) *Gaiella/Microlunatus* and (**D**) fungi/bacteria ratios. Legend: topographic wetness index (TWI), digital elevation model (DEM), length slope factor (LS factor), normalized difference water index (NDWI), lanthanum/ytterbium (La/Yb), and lanthanum/samarium olive fruit (La/SmO).

In the case of the *Gaiella/Rubrobacter* ratio (Figure 5B), latitude (Y) and longitude (X) were both identified as predictors. The OTU354/OTU2 ratio (*Gaiella/Rubrobacter*) prediction model showed a different pattern from the *Rubrobacter/Microlunatus* ratio model, with a high value associated to sandstone (upper part). From a geomorphological point of view, the model showed a hydrological pattern, a flow pathway, connected to the different spatial distribution of the bacterial *Gaiella/Rubrobacter* ratio values (Figure 5B). Indeed, TWI was a significant variable explaining the variations in *Gaiella/Rubrobacter* ratio (Figure 6B). The *Gaiella/Rubrobacter* ratio values ranged from 0.00046 to 0.076 in conglomeratic- and sandstone-derived soil, respectively. Conglomeratic and basaltic soils showed little to no presence of OTU354 (*Gaiella*). Significant variables explaining spatial variation in the *Gaiella/Microlunatus* ratio were elevation (DEM), La/Yb, and landforms. The low values of the *Gaiella/Microlunatus* ratio were mainly distributed in the central part of the basin and in the northeastern parts, strongly correlated with the topographic position and the lithology. Claystone showed lower *Gaiella/Microlunatus* ratio values than limestone or sandstone.

The fungi/bacteria ratio (Figure 5D) of the Greve Basin ranged from 1.89 to 10.64 in limestone- and fluvio-conglomeratic-derived soils, respectively (Table S4). The spatial distribution of the F:B ratio showed opposite range values between the eastern part and the western part of the basin. The eastern flank, dominated by conglomeratic and

alluvial deposits, showed the same pattern with higher values as the upper part of the basin, characterized by arenaceous outcrops. Interestingly, the F:B ratio model showed a different trend from the *Gaiella*/*Rubrobacter* and *Gaiella*/*Microlunatus* spatial prediction distribution maps (Figure 5). MLRA analyses showed that La/SmO, Lithology, NDWI, and TWI were the significant predictors of spatial variation in fungi/bacteria ratio (Figure 6D) for the Greve Basin.

4. Discussion

The aim of this work was to improve knowledge of how topography, together with lithology, shapes the composition and diversity of the soil microbial community. We hypothesized that the soils of ancient olive groves are inhabited by different microbial communities, and within them, diverse Actinobacteria are selected by the changing environmental and geoclimatic conditions. Stochastic modeling predicted significative OTUs belonging to the *Actinobacteria* genera. Specifically, OTU2 (*Rubrobacter*), OTU3 (*Microlunatus*), and OTU354 (*Gaiella*) were used to test our hypothesis because they were considered as microbial bioindicators for the Mediterranean environmental. A previous study conducted by the authors on the microbial composition of olive grove soils in the same area highlighted by decision tree analysis that *Rubrobacter* was a discriminant factor for clustering soils developed on carbonate flysch sediment and claystone. While *Microlunatus* in graded fluvio-lacustrine soils, *Gaiella* in graded arenaceous soils [2]. In the present research, we found positive correlations between altitude, slope aspect, topographic position, wetness indices, and bioindicators. Landscape position has been shown to be a driver of soil microbial community composition and functionality through its control over variations in water and nutrient flow, also under fire disturbance events [38,39].

Through the ML approach, elevation was found to be the most important predictive factor that showed a linear relationship with the presence and abundance of specific Actinobacteria (Figure 5). Previous work from forest sites located in the Italian Alps showed that altitude influences microbial structure and activity [40].

A random forest model revealed different spatial co-occurrence between OTUs belonging to Actinobacteria and a large number of geochemical factors, including trace elements (Li, Sb, Mo, Tl, V, La, Sm, and Yb—Figure 4); these geochemical factors were revealed to be more important in terms of explanatory variables for *Rubrobacter* and *Gaiella* than for *Microlunatus*. The differential geochemical driver for the biomarker spatial distribution may be coincident with specific weathering processes. In the case of Li, it is connected to silicate weathering [41]; indeed, the Li isotope is used as a proxy for marine carbonate sediments' paleo-weathering processes [42]. During the weathering of primary silicates, Li^+ can be adsorbed on the negatively charged of the clays or enter into the clay lattice, largely substituting Mg^{2+} and Fe^{2+} [43]. Furthermore, Li^+ can also be incorporated into ferromanganese oxides and oxyhydroxides and into the lattice of Mn^{2+} oxides found in biogenic rock varnish [44,45]. Interestingly, in this study, Li was found to be an important predictive factor for determining the spatial distribution of *Rubrobacter* through the watershed, especially associated to carbonatic flysch- and basaltic-derived soils (Figures 3B and 4); in fact, Li has been found relatively abundant in basalt-forming minerals [46].

Furthermore, we found that topography exerts a strong influence over the extremophile *Rubrobacter*, showing how southeast-facing slopes favor its spatial distribution compared to mesophilic *Microlunatus* (OTU2/3 model, Figure 5A). *Rubrobacter* is classified as a generalist in different types of ecosystems, i.e., it shows wide tolerance and distribution in many habitats, with a potential role in nitrogen as well as carbon metabolism [47].

For the Chianti study area, *Rubrobacter* is the most important genus in limestone and claystone soils and decreases in more arenaceous soils (Figure 5A). These latter soils,

developed on siliciclastic deposits, contain high levels of Fe (0.7 ppm) and a relatively high Mn content (13.3 ppm) with respect to the investigated lithologies. Furthermore, they are characterized by a low soil organic carbon content (SOC) equal to 0.54% and negative cerium (Ce) anomalies [17]. In these arenaceous soils, dominated by a sand fraction content (76%; Table 1) and mostly terraced, the OTUs belonging to the *Gaiella* genus were the most represented (Figure 5B,C). Other authors have related *Gaiella* to the absorption of phosphate and the decomposition of organic matter in a late stage of decomposition [48]. These arenaceous soils are characterized by a high hydraulic conductivity, avoiding the occurrence of ponding. The soil erosion is limited by the relatively flat slope of the terraces that reduce the occurrence of downslope surface runoff. However, in the inner part of the terraces, due to the relatively low slope and the tillage, an excess of overland flow runoff may occur due to soil saturation. These empirical evidences agree with those suggested by the spatial distribution models of *Gaiella* vs. *Rubrobacter* and *Gaiella* vs. *Microlunatus* (Figure 5B,C) and the RF models (Figure 4); in fact, *Gaiella* is found distributed mainly in the upper part of the slope, and it is correlated to the CND geomorphometric parameter (Figure 4) that represents the distance of the overland flow generation from the channel network. As suggested by Figure 5B, the prediction would lead to a higher presence of *Gaiella* than *Rubrobacter* in relation to the hydrological network (flow path). Soil erosion and deposition processes have been shown to affect soil microbial diversity and co-occurrence networks [49]. Although no forms of surface erosion have currently been detected, in the past, they may have led to the redistribution of different keystone taxa within the Greve Basin.

Previously, studies have found that Actinobacteria are affected by soil wetness conditions and negatively affected by soil moisture. In Southwestern Australia, under a Mediterranean climate, an increase in the relative abundance of Actinobacteria under water deficit conditions was found [50]. Indeed, water limitation can alter microbial community structure [51]. This corresponds to our finding where a positive correlation was found between the NDWI, TWI, and the fungi/bacteria ratio, indicating that these variables exert fundamental landscape-scale control over biogeochemical cycles [39] and microbial spatial distribution. Indeed, the F:B ratio is highly vulnerable to changing environmental conditions, such as climate and land use. Whereas it is not possible at present to predict soil microbial composition based on spectral reflectance, we can assert that measuring vegetation multispectral or hyperspectral reflectance could be helpful in detecting changes in microbial communities. In this regard, the La/Sm fractionation ratio related to the olive fruit was the most important variable to predict the spatial distribution of the F:B ratio (Figure 6D). Pelacani et al. (2024) showed that the soil signal or soil fingerprint in terms of the REE composition was also reported for the respective olive drupes for most elements. The La/Sm spatial distribution was also related to the NDVI value and La/YB fractionation pattern [17]. Since higher NDVI values are related to efficient photosynthesis activity and nutrient availability, a positive association between soil microbial biomass, its diversity, and plant physiological activity could be revealed. The values of the bioindicator F:B ratio were in the range of a previous study conducted in France [52]. Soil texture influenced the bioindicator F:B ratio; conglomerate soils showed a higher value on average than calcareous soils, showing a positive relationship with soil fungal biomass. Some studies have already previously created territorial maps regarding the distribution of the F:B ratio. It has been shown that among the characteristics of the soil, the pH, the organic carbon content, and the carbon/nitrogen ratio were the main factors influencing the distribution patterns of the F:B values [52]. We found that the REE fractionation ratio was the main driver for the biogeographic spatial distribution of the F:B value. In agreement with [53], the REEs are a powerful tool to trace pedogenesis in non-polluted soils deriving from different parent rock

materials. Diverse soil parameters influence REE uptake, such as the pH, Eh, CEC, OM, Fe-Mn-oxides, and soil textural fraction. Moreover, recycling of REEs by plants modifies their speciation, affecting the enrichment and fractionation of REEs in the soil–plant system.

Regarding pyroclastic soils, *Gemmatimonadetes* and *Acidobacteria* levels were significantly higher than those of *Rubrobacter*. This can be attributed to the chemical-physical characteristics of the soils, where soils of volcanic origin are characterized by a slightly acidic pH (6.28) and have a high sand fraction (80.6%), while soils dominated by *Rubrobacter* are characterized by a slightly-to-medium alkaline pH (7.5–8.2) and a 20–30% clay fraction (Table 1). Studies revealed that some species belonging to the phylum of *Actinobacteria* have the capacity of nitrogen-fixing [47]. On the other hand, nitrogen-fixing bacteria struggle in acidic soils, while the fungal population benefits in an acidic pH. Consistent with previous studies, the abundance of *Acidobacteria* and *Gemmatimonadetes* decreased under elevated N conditions [54].

Our results overall underline that spatial prediction maps of biomarkers, as demonstrated for *Actinobacteria* and the F:B ratio, can be generated starting from a combined geomorphometric and geochemical approach, given their importance as the main drivers of the soil microbial OTU distribution.

Overall, the approach used in the present research can make an important contribution to predicting soil physiological response to different climate change scenarios. *Rubrobacter*, *Acidobacteria*, and *Gemmatimonodates* are all considered "k-strategists", that is, they have slow growth rates and adaptation to utilize minimal resources [55]. On the other hand, *Rubrobacter* is considered a generalist in many environments, whereas *Acidobacteria* and *Gemmatimonodates* oligotrophic microbial taxa are more associated with soils with lower organic carbon availability [56]. The carbon/nitrogen (C/N) ratio in soil is considered an important indicator of C availability, where a high C/N ratio represents low C availability. Since it is known that individual microbial taxa exhibit different life strategies, the r/K ratio, like the fungi/bacteria ratio, could represent soil heterotrophic respiration [57]. Thus, in a future step of this research, the underlying mechanisms of the responses of soil heterotrophic respiration to drought will be explored by using the microbial life strategy approach.

5. Conclusions

This study is the first attempt to examine an ML approach to exploring the soil *Actinobacteria* biomarker variation in representative Mediterranean landforms, coupling different techniques. Our data support the hypothesis that geomorphometric factors predominantly shaped molecular biomarkers in the Chianti area, Tuscany, Italy. The RF model was a powerful tool to precisely analyze the association between a candidate biomarker and spatial processes, such as geomorphic and pedogenetic processes. The complex abiotic–biotic interactions were in part explained by the MLRA models that made it possible to identify the linear relationship between specific biomarkers, topographic factors, and REEs fractionation pattern. Both approaches explored different aspects of *Actinobacteria* distribution patterns and can be useful in ecology studies on soil management and function for sustainable agriculture. For future applications, the proposed approaches might need to be validated in other Mediterranean regions, considering that belonging to *Actinobacteria* was identified as a rare extremophile genus that could be revealed as an important functionality for land restoration. To this end, the use of multispectral or hyperspectral sensors and image processing techniques, given their scalable nature as well as their flexibility and low cost, could be useful to study the non-linear relation between soil microbial diversity and environmental factors.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/land14030583/s1, Figure S1: Relative abundance of Actinobacteria, Gemmatimonadetes, and Acidobacteria phylum in topsoil of 16 Tuscany olive groves; the letters in the abscissa axes refer to the sampled lithologies; for specifications, refer to Figure 2 of the article text; Table S1: Summary of the performance of the random forest model for three OTUs that are keystones for Mediterranean environments: OTU2, OTU3, and OTU354, belonging to *Actinobacteria*; Table S2: Summary of the performance of multiple regression analyses model for four bioindicators of Mediterranean environments; Table S3: Total REE contents in the DTPA-bioavailable topsoil fraction for contrasting lithologies of olive grove, Tuscany, Italy, and representative UCC-normalized REE fractionation patterns used in the models, denoted by subscript "n". Note: La/SmO and La/YbO refer to the olive; Table S4: Soil microbial bioindicators, medium values, concerning the investigated lithologies of olive grove topsoil located in the Chianti area (Tuscany, Italy). The letters of groups are referred to in Figure 2 in the article text.

Author Contributions: Conceptualization, S.P.; methodology, S.P., F.B. and M.T.C.; formal analysis, S.P. and F.B.; funding acquisition, S.P. and S.M.; investigation, S.P., M.T.C. and S.T.; project administration, S.M.; supervision, S.M. and S.T.; writing—original draft, S.P. and M.T.C.; writing—review and editing, S.M., S.T., S.P., M.T.C. and F.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Tuscany Region PSR 2014–2020—PS-GO 16.2, GeOEVO-App Project "A new decision support technology for authentication and geographic traceability of extra virgin olive oil" and Tuscany Region PSR 2014–2022 Mis. 16.2 GeOEVO Web GIS Project "A WEB GIS platform for the biogeochemical characterization of Tuscan olive groves" grant n. CUP ARTEA 1073640.

Data Availability Statement: Data developed in this study will be made available upon request to the corresponding author.

Acknowledgments: The authors are thankful to the farms belonging to "Frantoio Il Faggeto" (Anghiari, Arezzo, Italy) and "Frantoio del Greve Pesa" and "Frantoio Pruneti" (Florence, Italy) for providing olive orchards; to Angela Roccotelli for DNA extraction and quantification; and to Silvia Massagni and Melania Scacciati for the technical-bureaucratic support.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

OTU	Operational Taxonomic Unit
REE	Rare Earth Element
CNBL	Channel Network Base Level
VD	Valley Depth
CND	Vertical Distance to Channel Network
TRI	Terrain Ruggedness Index
TPI	Topographic Position Index
TWI	Topographic Wetness Index
DEM	Digital Elevation Model
LS Factor	Length Slope Factor
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index

References

1. Fausto, C.; Mininni, A.N.; Sofo, A.; Crecchio, C.; Scagliola, M.; Dichio, B.; Xiloyannis, C. Olive orchard microbiome: Characterisation of bacterial communities in soil-plant compartments and their comparison between sustainable and conventional soil management systems. *Plant Ecol. Divers.* **2018**, *11*, 597–610. [CrossRef]

2. Roccotelli, A.; Tommasini, S.; Ceccherini, M.T.; Calamai, L.; Ferrari, M.; Ghiotto, M.; Riccio, R.; Bonciani, L.; Pietramellara, G.; Moretti, S.; et al. Rare earth elements distribution and bacteriome to assess and characterize the soil landscapes of old olive orchards. *Diversity* **2024**, *16*, 427. [CrossRef]
3. Barka, E.A.; Vatsa, P.; Sanchez, L.; Gaveau-Vaillant, N.; Jacquard, C.; Klenk, H.-P.; Clément, C.; Ouhdouch, Y.; van Wezel, G.P. Taxonomy, physiology, and natural products of actinobacteria. *Microbiol. Mol. Biol. Rev.* **2016**, *80*, 1–43. [CrossRef] [PubMed]
4. Thirumurugan, D.; Vijayakumar, R.; Vadivalagan, C.; Karthika, P.; Khan MK, A. Isolation, structure elucidation and antibacterial activity of methyl-4,8-dimethylundecanate from the marine actinobacterium streptomyces albogriseolus ecr64. *Microb. Pathog.* **2018**, *121*, 166–172. [CrossRef]
5. Riquelme, C.; Marshall Hathaway, J.J.; Enes Dapkevicius, M.d.L.N.; Miller, A.Z.; Kooser, A.; Northup, D.E.; Jurado, V.; Fernandez, O.; Saiz-Jimenez, C.; Cheeptham, N. Actinobacterial diversity in volcanic caves and associated geomicrobiological interactions. *Front. Microbiol.* **2015**, *6*, 1342. [CrossRef]
6. Gousterova, A.; Paskaleva, D.; Vasileva-Tonkova, E. Characterization of culturable thermophilic actinobacteria from Livingston island, Antarctica. *Int. Res. J. Biol. Sci.* **2014**, *3*, 2278–3202.
7. Crits-Christoph, A.; Robinson, C.K.; Barnum, T.; Fricke, W.F.; Davila, A.F.; Jedynak, B.; McKay, C.P.; DiRuggiero, J. Colonization patterns of soil microbial communities in the atacama desert. *Microbiome* **2013**, *1*, 28. [CrossRef] [PubMed]
8. Hazarika, S.N.; Thakur, D. Actinobacteria. In *Beneficial Microbes in Agro-Ecology*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 443–476.
9. Büntgen, U.; Reinig, F.; Verstege, A.; Piermattei, A.; Kunz, M.; Krusic, P.; Slavin, P.; Štěpánek, P.; Torbenson, M.; del Castillo, E.M. Recent summer warming over the western mediterranean region is unprecedented since medieval times. *Glob. Planet. Change* **2024**, *232*, 104336. [CrossRef]
10. González-Ubierna, S.; de la Cruz, M.T.; Casermeiro, M.Á. Climate factors mediate soil respiration dynamics in mediterranean agricultural environments: An empirical approach. *Soil Res.* **2014**, *52*, 543–553. [CrossRef]
11. Labouyrie, M.; Ballabio, C.; Romero, F.; Panagos, P.; Jones, A.; Schmid, M.W.; Mikryukov, V.; Dulya, O.; Tedersoo, L.; Bahram, M. Patterns in soil microbial diversity across europe. *Nat. Commun.* **2023**, *14*, 3311. [CrossRef]
12. Dacal, M.; Garcia-Palacios, P.; Asensio, S.; Wang, J.; Singh, B.K.; Maestre, F.T. Climate change legacies contrastingly affect the resistance and resilience of soil microbial communities and multifunctionality to extreme drought. *Funct. Ecol.* **2022**, *36*, 908–920. [CrossRef]
13. Delgado-Baquerizo, M.; Bissett, A.; Eldridge, D.J.; Maestre, F.T.; He, J.-Z.; Wang, J.-T.; Hamonts, K.; Liu, Y.-R.; Singh, B.K.; Fierer, N. Palaeoclimate explains a unique proportion of the global variation in soil bacterial communities. *Nat. Ecol. Evol.* **2017**, *1*, 1339–1347. [CrossRef] [PubMed]
14. Morrissey, E.M.; Mau, R.L.; Hayer, M.; Liu, X.-J.A.; Schwartz, E.; Dijkstra, P.; Koch, B.J.; Allen, K.; Blazewicz, S.J.; Hofmockel, K. Evolutionary history constrains microbial traits across environmental variation. *Nat. Ecol. Evol.* **2019**, *3*, 1064–1069. [CrossRef]
15. De Vries, F.T.; Shade, A. Controls on soil microbial community stability under climate change. *Front. Microbiol.* **2013**, *4*, 265. [CrossRef]
16. Evans, S.E.; Wallenstein, M.D. Climate change alters ecological strategies of soil bacteria. *Ecol. Lett.* **2014**, *17*, 155–164. [CrossRef] [PubMed]
17. Pelacani, S.; Maerker, M.; Tommasini, S.; Moretti, S. Combining biodiversity and geodiversity on landscape scale: A novel approach using rare earth elements and spatial distribution models in an agricultural mediterranean landscape. *Ecol. Indic.* **2024**, *158*, 111583. [CrossRef]
18. Melloni, R.; Cardoso, E.J.B.N. Microbiome associated with olive cultivation: A review. *Plants* **2023**, *12*, 897. [CrossRef]
19. Sofo, A.; Ricciuti, P.; Fausto, C.; Mininni, A.N.; Crecchio, C.; Scagliola, M.; Malerba, A.D.; Xiloyannis, C.; Dichio, B. The metabolic and genetic diversity of soil bacterial communities depends on the soil management system and c/n dynamics: The case of sustainable and conventional olive groves. *Appl. Soil Ecol.* **2019**, *137*, 21–28. [CrossRef]
20. Karimi, B.; Terrat, S.; Dequiedt, S.; Saby NP, A.; Horrigue, W.; Lelièvre, M.; Nowak, V.; Jolivet, C.; Arrouays, D.; Wincker, P. Biogeography of soil bacteria and archaea across france. *Sci. Adv.* **2018**, *4*, eaat1808. [CrossRef]
21. Urbanová, M.; Šnajdr, J.; Baldrian, P. Composition of fungal and bacterial communities in forest litter and soil is largely determined by dominant trees. *Soil Biol. Biochem.* **2015**, *84*, 53–64. [CrossRef]
22. Ceccherini, M.T.; Ascher, J.; Pietramellara, G.; Mocali, S.; Viti, C.; Nannipieri, P. The effect of pharmaceutical waste-fungal biomass, treated to degrade dna, on the composition of eubacterial and ammonia oxidizing populations of soil. *Biol. Fertil. Soils* **2007**, *44*, 299–306. [CrossRef]
23. Santoni, M.; Verdi, L.; Imran Pathan, S.; Napoli, M.; Dalla Marta, A.; Dani, F.R.; Pacini, G.C.; Ceccherini, M.T. Soil microbiome biomass, activity, composition and CO_2 emissions in a long-term organic and conventional farming systems. *Soil Use Manag.* **2023**, *39*, 588–605. [CrossRef]
24. Bokulich, N.A.; Subramanian, S.; Faith, J.J.; Gevers, D.; Gordon, J.I.; Knight, R.; Mills, D.A.; Caporaso, J.G. Quality-filtering vastly improves diversity estimates from illumina amplicon sequencing. *Nat. Methods* **2013**, *10*, 57–59. [CrossRef] [PubMed]

25. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I. Qiime allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335–336. [CrossRef] [PubMed]
26. Haas, B.J.; Gevers, D.; Earl, A.M.; Feldgarden, M.; Ward, D.V.; Giannoukos, G.; Ciulla, D.; Tabbaa, D.; Highlander, S.K.; Sodergren, E. Chimeric 16s rrna sequence formation and detection in sanger and 454-pyrosequenced pcr amplicons. *Genome Res.* **2011**, *21*, 494–504. [CrossRef]
27. Edgar, R.C.; Haas, B.J.; Clemente, J.C.; Quince, C.; Knight, R. Uchime improves sensitivity and speed of chimera detection. *Bioinformatics* **2011**, *27*, 2194–2200. [CrossRef]
28. Edgar, R.C. Uparse: Highly accurate otu sequences from microbial amplicon reads. *Nat. Methods* **2013**, *10*, 996–998. [CrossRef]
29. Wang, Q.; Garrity, G.M.; Tiedje, J.M.; Cole, J.R. Naive bayesian classifier for rapid assignment of rrna sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **2007**, *73*, 5261–5267. [CrossRef]
30. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The silva ribosomal rna gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2012**, *41*, D590–D596. [CrossRef]
31. Chao, A.; Bunge, J. Estimating the number of species in a stochastic abundance model. *Biometrics* **2002**, *58*, 531–539. [CrossRef]
32. Wang, L.; Liu, H. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. *Int. J. Geogr. Inf. Sci.* **2006**, *20*, 193–213. [CrossRef]
33. Sörensen, R.; Zinko, U.; Seibert, J. On the calculation of the topographic wetness index: Evaluation of different methods based on field observations. *Hydrol. Earth Syst. Sci.* **2006**, *10*, 101–112. [CrossRef]
34. Conrad, O.; Bechtel, B.; Bock, M.; Dietrich, H.; Fischer, E.; Gerlitz, L.; Wehberg, J.; Wichmann, V.; Böhner, J. System for automated geoscientific analyses (saga) v. 2.1. 4. *Geosci. Model Dev.* **2015**, *8*, 1991–2007. [CrossRef]
35. Guisan, A.; Weiss, S.B.; Weiss, A.D. Glm versus cca spatial modeling of plant species distribution. *Plant Ecol.* **1999**, *143*, 107–122. [CrossRef]
36. Riley, S.J.; DeGloria, S.D.; Elliot, R. Index that quantifies topographic heterogeneity. *Intermt. J. Sci.* **1999**, *5*, 23–27.
37. Bahrenberg, G.; Giese, E.; Nipper, J. Statistische Methoden in der Geographie. In *Multivariate Statistik 2*; Gebr. Borntraeger: Berlin, Germany, 1992; Volume 2.
38. Brockett BF, T.; Prescott, C.E.; Grayston, S.J. Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western canada. *Soil Biol. Biochem.* **2012**, *44*, 9–20. [CrossRef]
39. Fairbanks, D.; Shepard, C.; Murphy, M.; Rasmussen, C.; Chorover, J.; Rich, V.; Gallery, R. Depth and topographic controls on microbial activity in a recently burned sub-alpine catchment. *Soil Biol. Biochem.* **2020**, *148*, 107844. [CrossRef]
40. Siles, J.A.; Cajthaml, T.; Minerbi, S.; Margesin, R. Effect of altitude and season on microbial activity, abundance and community structure in alpine forest soils. *FEMS Microbiol. Ecol.* **2016**, *92*, fiw008. [CrossRef]
41. Day, C.C.; von Strandmann, P.A.E.P.; Mason, A.J. Lithium isotopes and partition coefficients in inorganic carbonates: Proxy calibration for weathering reconstruction. *Geochim. Et Cosmochim. Acta* **2021**, *305*, 243–262. [CrossRef]
42. von Strandmann, P.A.E.P.; Dellinger, M.; West, A.J. *Lithium Isotopes: A Tracer of Past and Present Silicate Weathering*; Cambridge University Press: Cambridge, UK, 2021.
43. Vigier, N.; Decarreau, A.; Millot, R.; Carignan, J.; Petit, S.; France-Lanord, C. Quantifying li isotope fractionation during smectite formation and implications for the li cycle. *Geochim. Et Cosmochim. Acta* **2008**, *72*, 780–792. [CrossRef]
44. Jiang, X.; Lin, X.; Yao, D.; Zhai, S.; Guo, W. Geochemistry of lithium in marine ferromanganese oxide deposits. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **2007**, *54*, 85–98. [CrossRef]
45. Yeager, C.M. Life on the Edge: Microbes in Rock Varnish. October 2019. Available online: https://doi.org/10.2172/1570605 (accessed on 15 January 2025).
46. Fairén, A.G.; Losa-Adams, E.; Gil-Lozano, C.; Gago-Duport, L.; Uceda, E.R.; Squyres, S.W.; Rodríguez, J.A.P.; Davila, A.F.; McKay, C.P. Tracking the weathering of basalts on mars using lithium isotope fractionation models. *Geochem. Geophys. Geosystems* **2015**, *16*, 1172–1197. [CrossRef] [PubMed]
47. Zhang, B.; Wu, X.; Tai, X.; Sun, L.; Wu, M.; Zhang, W.; Chen, X.; Zhang, G.; Chen, T.; Liu, G.; et al. Variation in actinobacterial community composition and potential function in different soil ecosystems belonging to the arid heihe river basin of northwest china. *Front. Microbiol.* **2019**, *10*, 2209. [CrossRef]
48. Severino, R.; Froufe, H.J.C.; Barroso, C.; Albuquerque, L.; Lobo-da-Cunha, A.; da Costa, M.S.; Egas, C. High-quality draft genome sequence of gaiella occulta isolated from a 150 meter deep mineral water borehole and comparison with the genome sequences of other deep-branching lineages of the phylum actinobacteria. *MicrobiologyOpen* **2019**, *8*, e00840. [CrossRef]
49. Yang, Q.; Peng, J.; Ni, S.; Zhang, C.; Wang, J.; Cai, C. Erosion and deposition significantly affect the microbial diversity, co-occurrence network, and multifunctionality in agricultural soils of northeast china. *J. Soils Sediments* **2024**, *24*, 888–900. [CrossRef]
50. Mickan, B.S.; Abbott, L.K.; Solaiman, Z.M.; Mathes, F.; Siddique, K.H.M.; Jenkins, S.N. Soil disturbance and water stress interact to influence arbuscular mycorrhizal fungi, rhizosphere bacteria and potential for n and c cycling in an agricultural soil. *Biol. Fertil. Soils* **2019**, *55*, 53–66. [CrossRef]

51. Jaeger, A.C.H.; Hartmann, M.; Conz, R.F.; Six, J.; Solly, E.F. Prolonged water limitation shifts the soil microbiome from copiotrophic to oligotrophic lifestyles in scots pine mesocosms. *Environ. Microbiol. Rep.* **2024**, *16*, e13211. [CrossRef]
52. Djemiel, C.; Dequiedt, S.; Bailly, A.; Tripied, J.; Lelièvre, M.; Horrigue, W.; Jolivet, C.; Bispo, A.; Saby, N.; Valé, M.; et al. Biogeographical patterns of the soil fungal:Bacterial ratio across france. *mSphere* **2023**, *8*, e00365-23. [CrossRef]
53. Laveuf, C.; Cornu, S. A review on the potentiality of Rare Earth Elements to trace pedogenetic processes. *Geoderma* **2009**, *154*, 1–9. [CrossRef]
54. Cederlund, H.; Wessén, E.; Enwall, K.; Jones, C.M.; Juhanson, J.; Pell, M.; Philippot, L.; Hallin, S. Soil carbon quality and nitrogen fertilization structure bacterial communities with predictable responses of major bacterial phyla. *Appl. Soil Ecol.* **2014**, *84*, 62–68. [CrossRef]
55. Davis, K.E.R.; Sangwan, P.; Janssen, P.H. Acidobacteria, rubrobacteridae and chloroflexi are abundant among very slow-growing and mini-colony-forming soil bacteria. *Environ. Microbiol.* **2011**, *13*, 798–805. [CrossRef] [PubMed]
56. Fierer, N.; Bradford, M.A.; Jackson, R.B. Toward an ecological classification of soil bacteria. *Ecology* **2007**, *88*, 1354–1364. [CrossRef] [PubMed]
57. Chen, H.; Jing, Q.; Liu, X.; Zhou, X.; Fang, C.; Li, B.; Zhou, S.; Nie, M. Microbial respiratory thermal adaptation is regulated by r-/k-strategy dominance. *Ecol. Lett.* **2022**, *25*, 2489–2499. [CrossRef] [PubMed]