

RESEARCH

Open Access

# Microbiome of vineyard soils is shaped by geography and management



Emanuela Coller<sup>1†</sup>, Alessandro Cestaro<sup>2\*\*†</sup> , Roberto Zanzotti<sup>3</sup>, Daniela Bertoldi<sup>3</sup>, Massimo Pindo<sup>2</sup>, Simone Larger<sup>2</sup>, Davide Albanese<sup>2</sup>, Enzo Mescalchin<sup>3</sup> and Claudio Donati<sup>2</sup>

## Abstract

**Background:** Despite their importance as a reservoir of biodiversity, the factors shaping soil microbial communities and the extent by which these are impacted by cultivation are still poorly understood. Using 16S rRNA gene and ITS sequencing, we characterized the soil microbiota of vineyards and of neighboring permanent grassland soils in the Italian province of Trentino, and correlated their structure and composition to location, chemical properties of the soil, and land management.

**Results:** Bacterial communities had a core of conserved taxa accounting for more than 60% of the reads of each sample, that was influenced both by geography and cultivation. The core fungal microbiota was much smaller and dominated by geography alone. Cultivation altered the structure and composition of the soil microbiota both for bacteria and fungi, with site-specific effects on their diversity. The diversity of bacterial and fungal communities was generally inversely correlated across locations. We identified several taxa that were impacted by the chemical properties and texture of the soil.

**Conclusions:** Our results highlight the different responses of bacterial and fungal communities to environmental factors and highlight the need to characterize both components of the soil microbiota to fully understand the factors that drive their variability.

## Background

The microorganisms that colonize soil are amongst the most abundant and diverse life forms on Earth, contributing to all geochemical processes on a global scale [1], and constituting a rich source of yet uncharacterized natural products of potential interest for pharmaceutical or biotechnological applications [2]. The biodiversity of soil microbial communities is increasingly recognized as a major factor for human health both directly, by limiting the spread of potential pathogens, and indirectly, by contributing to processes that provide clean air, water, and healthy food [3]. Soil serves as a primary reservoir for plant-colonizing bacteria [4], that play a major role in determining plant productivity [5] and preventing invasion by bacterial pathogens [6].

Bacteria, archaea, and fungi are the dominant components of soil microbiota, generally accounting for more than 99% of the microbial biomass in soil samples [7]. Large-scale surveys have shown that the diversity encompassed by soil microbial communities exceeds what is found in host-associated communities [8, 9], probably as a consequence of the enormous range of environmental conditions that can be experienced by microorganisms in surface soils [7, 10]. Bacterial communities are characterized by pronounced heterogeneity at small spatial scales, and by a more homogeneous structure over large spatial scales [10], showing biogeographical patterns that are significantly weaker than what is found for plants [11]. Besides bacteria, fungi are the other major component of the soil microbiota, playing crucial roles both as saprotrophs, plant mutualists, and pathogens [8] and competing with bacteria for access to nutrients through the production of antimicrobial compounds [12]. Large-scale studies have shown that a large fraction of fungal taxa found in soils is not represented in sequence databases, that the diversity of soil fungal communities is influenced by a variety of

\* Correspondence: [alessandro.cestaro@fmach.it](mailto:alessandro.cestaro@fmach.it)

<sup>†</sup>Emanuela Coller and Alessandro Cestaro contributed equally to this work.

<sup>2</sup>Unit of Computational Biology, Research and Innovation Centre, Fondazione Edmund Mach, Via Mach 1, 38010 San Michele all'Adige, Italy  
Full list of author information is available at the end of the article



climatic and edaphic factors [13, 14], and that soil fungal communities exhibit evident patterns of geographical clustering [8].

Understanding how agricultural practices impact the soil microbiota is an important subject towards a more sustainable agriculture. Recent studies on deforested lands have shown that land use has long-term effects on soil microbiota structure and diversity [15]. Both parameters are consistently altered by high levels of nutrient inputs related to human activities [16]. However, it is only by comparing cultivated and non-cultivated soils across locations, that it is possible to quantify the relative weights of differences related to location compared to these related to cultivation. In this work, we have characterized the bacterial and fungal microbiota in soils collected in 10 sites from Trentino, a region in the Italian Alps, with the aim of defining the taxonomic structure of both the bacterial and fungal components of the soil microbiota, and study the relative effects of location, chemical characteristics of the soil, and land use. To this purpose, sites were chosen to have cultivated patches (vineyards) surrounded by permanent grasslands. In each site, samples were collected from the vineyards and from grasslands at different distances from the vineyards. By comparing cultivated and permanent grasslands from the same site, we identified the species that were consistently impacted by cultivation.

It has recently been shown that a relatively small number of ubiquitous species dominate the global soil microbiota [17], while, at the other end of the spectrum, rare taxa play an important role as a reservoir of biological functions and resiliency against environmental changes [18]. However, how the set of ubiquitous species changes from global to local scale, and if geographical factors or land use can significantly alter its size is not known. Moreover, despite their importance as a component of the soil microbiota, no data are available concerning the existence of a core of shared fungal species in soil samples and how this core depends on soil type and is impacted by human intervention. If similar general ecological mechanisms drive the establishment of bacterial and fungal microbiota in soil, we should expect that the size of the core of bacterial and fungal microbiota has the same dependency on location and land use. On the other hand, differences in the way land use or location impact the core bacterial and fungal microbiomes could highlight fundamental differences in the factors that drive the colonization of soil by these two classes of microorganisms.

Defining the species that characterize the soil of a given area and how these are influenced by external factors is especially relevant in the case of grapevine, since soil has been suggested to be the major source of grapevine-associated microbiota [4]. Although the existence of a connection between soil and grape microbiota is still debated [19], the correlation between grape microbiota and

wine metabolite profiles has suggested that microbial communities contribute to define the regional characteristics of wine [20, 21], leading to the hypothesis of the existence of a “microbial terroir” of wine grapes [21]. Here, we defined the core of ubiquitous bacterial and fungal species that were present in all soil samples and tested if its size was altered by location or type of cultivation. We found that while the bacterial component of the microbiome had a core of conserved species that accounted for more than 60% of the sequenced reads, and that was shaped both by location and land use, the core fungal microbiome was smaller and determined by geographic factors that dominated differences due to land management.

Finally, having characterized fungal and bacterial communities in the same samples allowed us to highlight the effects of their interactions. Recently, a global study of the variability of the topsoil microbiome has provided evidence of a strong antagonism between the two communities [12] mediated by antibiotic production by the soil fungi. Here, by correlating fungal and bacterial diversity in the same samples, we observed patterns consistent with the hypothesis that complex fungal communities impose a strong selection on bacterial community causing a loss of diversity in the latter, as expected in the case of antibiotic-mediated interactions [22].

## Materials and methods

### Sample collection

The sampling sites were identified in 10 vineyards from 4 different locations (Ala, Besagno, Mori and S. Felice) because of their contiguity, at least along 20 m, to perennial crop-covered surfaces (Additional file 1: Figure S1). The experimental protocol set 3 samplings points respectively between the rows (V) and in the perennial crop area at a distance of 8 (P1) and 16 (P2) meters from the border of the vineyard (see Additional file 2: Figure S2). The choice for 8 and 16 m guaranteed the reproducibility of sampling in each location, since the grassland areas surrounding vineyards have different and relatively small size. In all the sampling spots, grasses were mainly constituted from species of *Poaceae* family. In the vineyard sites (V), no predominant species were found while in the two grasslands (P1 and P2), the dominant species were *Arrhenatherum elatius*, *Bromus erectus*, and *Trisetum flavescens*. For each position 6 equally spaced repetitions were performed, for a total of 180 samples. Additional file 14: Table S1 shows site localization, sampling dates, and technical characteristics of the vineyards (planting year, previous crop). All samples had a similar range of soil texture (loam, sandy clay loam, sandy loam, and silty loam). Vineyards soils (V) were statistically different from permanent grassland (both P1 and P2) for lower quantity of soil organic matter (SOM) ( $50.63 \pm 1.43$  g/kg,  $70.17 \pm 2.32$  g/kg, and  $73.04 \pm 2.26$  g/kg respectively) and total nitrogen (N tot) ( $2.78 \pm$

0.07 g/kg (V),  $3.92 \pm 0.14$  g/kg (P1),  $4.07 \pm 0.13$  g/kg (P2)), instead they had higher concentrations of total carbonate (CaCO<sub>3</sub>) ( $199.73 \pm 14.82$  g/kg (V),  $171.35 \pm 15.01$  g/kg (P1),  $156.33 \pm 14.67$  g/kg (P2)) and available heavy metals (Cu DTPA  $51.23 \pm 3.94$  mg/kg (V),  $16.9 \pm 1.71$  mg/kg (P1),  $6.66 \pm 0.46$  mg/kg (P2), and Zn DTPA  $10.53 \pm 0.94$  mg/kg (V),  $7.7 \pm 0.60$  mg/kg (P1),  $6.21 \pm 0.57$  mg/kg (P2)). See Additional file 3: Figure S3.

Samplings were executed collecting 20 cm of soil by means of a manual, one-piece, 7-cm-diameter drill for loamy soils (Eijkkelkamp, Edelman model). For chemical analysis and for taxonomic purposes of bulk soil, the first 5 cm of soil were removed. Each sample consisted of 4 drillings that were homogenized in a signed, plastic bag. From every one of them, a small volume of soil was collected in a 50-ml tube and chilled to 6/8 °C during the sampling time after which they were frozen at -18 °C.

#### DNA extraction, library preparation, and sequencing

The soil samples were freeze-dried and sieved with a 0.2-mm-mesh size and stored at -80 °C until DNA extraction. Total DNA was extracted from 0.25 g of each composite soil sample using the PowerSoil DNA isolation kit (MO BIO Laboratories Inc., CA, USA) according to the manufacturer's instructions. Total genomic DNA was amplified using primers specific to either the bacterial and archaeal 16S rRNA gene or the fungal ITS1 region. The specific bacterial primer set 515F (5'-GTGYCAGCMGCCGCGGTAA-3') and the 806R (5'-GGACTACN VGGGTWTCTAAT-3') was used [23] with degenerate bases suggested by Apprill et al. [24] and by Parada et al. [25]. Although no approach based on PCR amplification is free from bias, this primer pair has been shown to guarantee good coverage of known bacterial and archaeal taxa [26]. For the identification of fungi, the internal transcribed spacer 1 (ITS1) was amplified using the primer ITS1F (5'-CTTGGTCATTTAGAGGAA GTAA-3') [27] and ITS2 (5'-GCTGCGTTCTTCATCGAT GC-3') [28]. All the primers included the specific overhang Illumina adapters for the amplicon library construction.

For the 16S V4 region, each sample was amplified by PCR using 25 µl reaction with 1 µM of each primer. More in detail 12.5 µl of 2× KAPA HiFi HotStart ReadyMix and 10 µl forward and reverse primers, were used in combination with 2.5 µl of template DNA (5–20 ng/µl). PCR reactions were executed by GeneAmp PCR System 9700 (Thermo Fisher Scientific) and the following cycling conditions: initial denaturation step at 95 °C for 5 min (1 cycle); 28 cycles at 95 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s; final extension step at 72 °C for 5 min (1 cycle).

For the ITS1 region, each sample was amplified by PCR using 25 µl reaction with 10 µM of each primer. More in detail 22 µl of premix FastStart High Fidelity PCR System (Roche) and 2 µl forward and reverse primers were used in combination with 1 µl of template DNA (5–20 ng/ul).

PCR reactions were executed by GeneAmp PCR System 9700 (Thermo Fisher Scientific) and the following cycling conditions: initial denaturation step at 95 °C for 3 min (1 cycle); 30 cycles at 95 °C for 20 s, 50 °C for 45 s, 72 °C for 90 s; final extension step at 72 °C for 10 min (1 cycle).

The amplification products were checked on 1.5% agarose gel and purified using the Agencourt AMPure XP system (Beckman Coulter, Brea, CA, USA), following the manufacturer's instructions. Afterward, a second PCR was used to apply dual indices and Illumina sequencing adapters Nextera XT Index Primer (Illumina), by 7 cycles PCR (16S Metagenomic Sequencing Library Preparation, Illumina). The amplicon libraries were purified using Agencourt AMPure XP system (Beckman), and the quality control was performed on a Typestation 2200 platform (Agilent Technologies, Santa Clara, CA, USA). Finally, all barcoded libraries were pooled in an equimolar way and sequenced on an Illumina® MiSeq (PE300) platform (MiSeq Control Software 2.5.0.5 and Real-Time Analysis software 1.18.54.0).

#### Bioinformatic processing of the sequences

The sequences were assigned to samples using sample-specific barcodes and saved in FASTQ-formatted files. Sequences were deposited to the European Nucleotide Archive (ENA) with study accession PRJEB31356. Raw data FASTQ files were analyzed using the software pipeline MICCA [29] v. 1.6.1 (Microbial Community Analysis).

Raw overlapping 16S paired-end reads were assembled (merged) using the procedure described in [30]. Paired-end reads with an overlap length smaller than 200 bp and with more than 50 mismatches were discarded. After trimming forward and reverse primers, merged reads shorter than 240 bp and with an expected error rate higher than 0.5% were removed.

Reads with less than 60% similarity to the sequences present in the Greengenes [31] database (clustered at 85%, release 13\_5) were discarded using VSEARCH [32] v2.3.4. Filtered sequences were clustered into operational taxonomic units (OTUs) at 97% identity using the denovo greedy algorithm available in MICCA. OTUs were taxonomically classified using the Ribosomal Database Project (RDP) Classifier [33] v2.11. Multiple sequence alignment (MSA) was performed on the denoised reads applying the Nearest Alignment Space Termination [29, 34] (NAST) algorithm and the phylogenetic tree was inferred using FastTree [35] (v2.1.8).

Raw overlapping ITS paired-end reads were merged, and merged sequences with an overlap length smaller than 100 bp and with more than 32 mismatches were discarded. After primer trimming, merged reads shorter than 150 bp and with an expected error rate higher than 0.5% were removed. Filtered sequences were clustered at 97% identity using the denovo greedy algorithm and OTUs

were taxonomically classified using the RDP Classifier v2.11 and the UNITE database (release 07/04/2014) [36]. Sequences that were not in the UNITE database were indicated as “unclassified” while sequences that are present, but without taxonomic information available, were classified as “fungi\_unidentified” plus a numeric suffix for the taxonomic level. For instance, in the case of the family level, sequences that have no taxonomic information other than being fungi were labeled as “fungi\_unidentified\_1\_1”.

To compensate for different sequencing depths, samples were rarefied to an even depth of 16,000 reads for 16S and 22,000 for ITS sequences.

### Statistical analysis of the data

Biom files were imported into R v3.4.3 using the *phyloseq* package [37] v1.22.3 for downstream statistical analysis. Alpha diversity was calculated using the Shannon entropy [38]. Beta-diversity was calculated using the Bray-Curtis distance. Permutational MANOVA (PERMANOVA) statistical tests were performed using the R package *vegan* v2.4–6 with the *adonis2()* function with 999 permutations. Taxa significantly different between vineyard soils and permanent grasslands were identified using the generalized linear models implemented in the R package *DESeq2* [39] v1.18.1 on the unrarified reads, filtered selecting OTUs that were represented by at least 10 reads in more than 25% of the samples to increase the reproducibility of the results [40]. Random Forest models to identify relevant environmental factors to predict alpha-diversity were built using the *randomForest* [41] v4.6–14 R package. Linear models of the alpha-diversity as a function of the environmental variables were built using the *lm* function in R, and contrasts between different locations and soil types were determined using the function *contrast* from the R package *contrast* v0.21. Multilevel hierarchical linear models were built using the *lmer* function of the *lme4* v1.1–17 R package, modeling slope as fixed effect and intercept as a random variable. *p* values were estimated using the package *lmerTest* 3.0–1. Correlation between taxa at all taxonomic levels and environmental variables were identified using the *MICtools* software package [42], a recent software package that identifies significant correlations between large datasets using an approach based on mutual information. Briefly, significant correlations were identified calculating the total information coefficient estimator  $TIC_e$  [43]. The associated (FDR corrected) *p* values were estimated using a permutation-based strategy. The strength of the association was then quantified using the  $MIC_e$  estimator of the maximal information coefficient [43] and the Spearman’s rank correlation coefficient  $\rho$ . Cladograms of the taxonomy were drawn using the R package *Metacoder* [44]. The size of the nodes was

proportional to the relative abundance of the taxa while the color represented Spearman’s rank correlation coefficient  $\rho$ .

### Texture and chemical characteristics of the soil

All samples were air-dried at room temperature and sieved < 2 mm. For the analysis of organic carbon and total inorganic carbon, 50 g of sieved soil were ground using an agate-ball mill (PM 4000, Retsch GMBH, Haan, Germany) to reduce the particle size < 0.02 mm.

Soil texture was determined as the percentage of sand (2.0–0.050 mm), silt (0.005–0.002 mm) and clay (< 0.002 mm) by wet sieving and with the use of hydrometer after dispersion with sodium hexametaphosphate.

Soil pH was measured in aqueous solution suspension (ratio 1: 2.5, soil: water) using an INOLAB LEVEL 2 (WTW, Weilheim, Germany) pH-meter equipped with a SenTixTM41 pH Electrode (WTW, Weilheim, Germany).

Total inorganic carbon was determined by volumetric method with a Dietrich-Fruehling calcimeter by measuring the CO<sub>2</sub> evolved after HCl treatment of soil according to ISO 10693 (ISO 10693:1995. Soil quality–determination of carbonate content–Volumetric method), whereas active lime was quantified by titration after reaction of soil with ammonium oxalate 0.1 M (Drouineau method). Both were expressed as g CaCO<sub>3</sub>/kg of soil.

Total C and N were measured after dry combustion in excess oxygen using a CN analyzer (MacroVario, Elementar, Langenselbold, Germany). Two hundred milligrams of milled soil were weighted in tin foil and analyzed following manufacturer’s instructions. Organic C was then calculated as the difference between total C and total inorganic C as reported by ISO 10694 (ISO 10694:19959. Soil quality–determination of organic and total carbon after dry combustion). The C/N ratio was calculated as the ratio between organic C and total N.

The available fractions of Cu, Zn, Pb, and Cd were extracted with a DTPA 0.005 M, CaCl<sub>2</sub> 0.01 M, and triethanolamine 0.1 M solution and measured with an ICPOES spectrometer (Optima 8300, Perkin Elmer, Waltham, USA) equipped with a cyclonic nebulizer, using the following wavelengths: Cu = 324.752 nm, Zn = 213.857 nm, Pb = 220.353 nm, Cd = 226.502 nm. For the quantification, the instrument was calibrated using a certified standard solution (Merck, Darmstadt, Germany).

### Results

Soil samples were collected from 10 sites in the Trentino province (Additional file 1: Figure S1 and Additional file 14: Table S1a). Each site was characterized by sampling the vineyard (V) and the grasslands at 8 (P1) and 16 m (P2) from the grapevine row (see “Methods” section). For each triplet, 6 samples were collected, for a total number of 180 samples. After discarding 3 samples

due to the low quantity of extracted DNA, a total of 5,705,432 amplicon sequences from the V4 region of 16S rRNA gene and 7,350,959 sequences from the ITS region were obtained from the remaining 177 samples. The 16S and ITS sequences were clustered into 21,113 and 12,542 operational taxonomic units (OTUs, 97% identity), respectively. After dropping one 16S sample due to the low number of reads, 16S and ITS samples were evenly rarefied to 16,000 and 22,000 reads per sample, respectively. After rarefaction, the dataset was composed of 19,584 bacterial and 12,101 fungal OTUs.

#### Microbiota composition is conditioned by cultivation for both bacterial and fungal communities

The dominant bacterial Phyla were *Acidobacteria* (22.7%), *Proteobacteria* (18.8%), and *Actinobacteria* (16.5%), while on average, 14.1% of the reads could not be classified. At the family level (Fig. 1a, c), the fraction of bacterial OTUs that could not be classified grew to 34%, while for those that were classified the dominant family was *Gp6* (13%), followed by *Nitrososphaeraceae* and *Planctomycetaceae* (9% and 5%, respectively). For fungi, the dominant Phyla were *Ascomycota* (51.8%), *Zygomycota* (20.1%), and *Basidiomycota* (11.2%), while 12.1% was constituted by unclassified OTUs. At the family level (Fig. 1b, d), the dominant taxa were *Mortierellaceae* (17.4%) followed by *Nectriaceae* (8.8%) and a family of unidentified *Ascomycota* (5.2%). The fraction of fungal OTUs not classified at the family level was 25%.

To identify taxa that were significantly impacted by cultivation, we modeled the read counts using generalized linear models [39] (GLMs) taking baseline differences between the sites into account. For this analysis, samples from both types of permanent grassland (both P1 and P2) were considered together. For bacteria, we found 336 OTUs that were significantly differentially abundant ( $p < 0.01$ , FDR corrected) between samples from vineyard soils and from permanent grasslands from the same site. Of these, 224 were higher in vineyard samples and 112 in grasslands (Additional file 15: Table S2). The 10 most significant ones included taxa from the families *Gp4*, *GP6*, *Hyphomicrobiaceae*, and *Planctomycetaceae* (Additional file 4: Figure S4). For fungi, we found 57 taxa significantly ( $p < 0.01$ , FDR corrected) more abundant in the vineyard than permanent grassland soil from the same site, and 37 taxa significantly less abundant in vineyard samples (Additional file 16: Table S3). Amongst these, the most significant were taxa from the family *Amphisphaeriaceae*, that in vineyard samples accounted for up to 20% of the fungal microbiota (Additional file 5: Figure S5). Other differently distributed taxa included unidentified *Pleosporales*, unidentified *Ascomycota*, *Hyaloscyphaceae*, and *Sordariomycetes*.

#### Bacterial communities have a core of conserved species that is shaped by both location and land use

The distribution of OTUs across samples (Fig. 2a) showed that the largest fraction of OTUs was specific to a small number of samples while only a small set of OTUs was ubiquitous. Specifically, out of 19,584 OTUs, only 162 were present in all samples, 484 were present in at least 80% of the samples, and 961 in at least 50% of the samples.

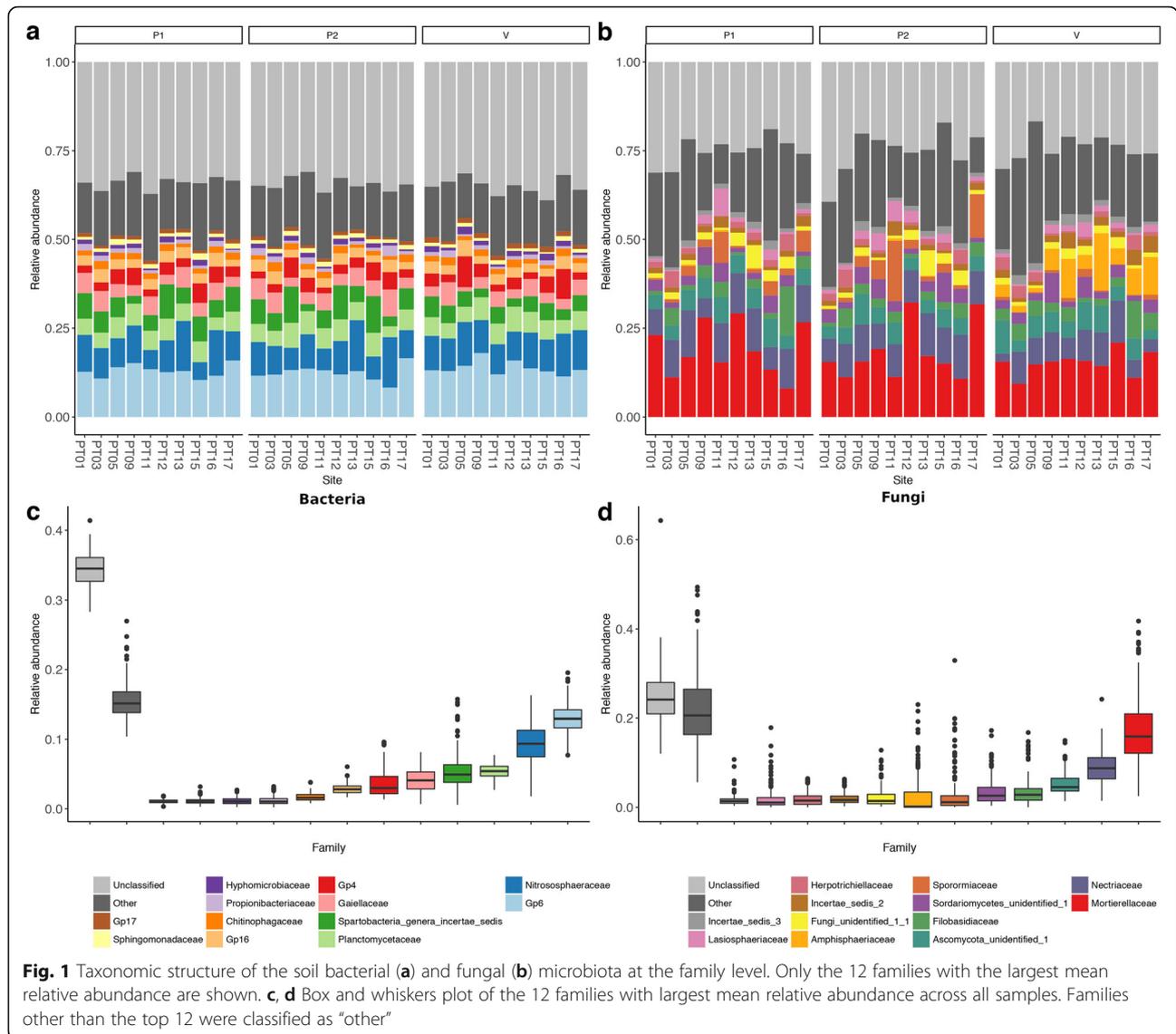
Despite the relatively low number, the core OTUs accounted on average for  $48 \pm 3.8\%$  of the bacterial reads of each sample. Using a more relaxed definition of the core OTUs, i.e., defining as “core” those OTUs that are present in more than 95% of the samples, core OTUs accounted on average for  $64 \pm 3\%$  of each microbiome (Fig. 2b). For comparison, the specific OTUs (i.e., those present in less than 5% of the samples) that are the vast majority of the bacterial OTUs, only accounted for an average of  $1.5 \pm 0.5\%$  of each bacterial microbiome. *Acidobacteria*, a minority component of the specific microbiome, was the dominant component of the core microbiome. Other frequent components of the core microbiome included *Actinobacteria* and *Proteobacteria*, while *Bacteroidetes* and *Firmicutes* were almost absent.

We next asked if the size of the core microbiome was influenced by the cultivation type or geographical origin of the samples. We found that the core microbiome of P1, P2, and V samples included 250, 256, and 240 OTUs, respectively, while the site-specific core microbiome varied between 372 and 638 OTUs (Additional file 17: Table S4). In both cases, the core microbiome was significantly larger than expected by random sampling (Wilcoxon rank-sum test,  $p$  values  $1.16 \times 10^{-6}$  for site and  $7.8 \times 10^{-3}$  for cultivation type, respectively), suggesting that both these factors favor the colonization of soil by a defined set of bacterial species.

In general, in any given dataset, the number of core OTUs (core microbiome) decreases with the number of samples, since each newly added sample might miss OTUs that were core in the reduced dataset. Therefore, it is possible to estimate the true size of the core microbiome, i.e., the number of OTUs that are always present in this kind of soils independently of the sampling size by extrapolating from a random subsampling. The results are shown in Additional file 6: Figure S6a, where we plot the number of core OTUs as a function of the number of samples. The curve can be fitted by a power-law decay converging to a plateau of  $117 \pm 3.5$  OTUs that is the estimated size of the core bacterial microbiome of these soils.

#### The core of conserved fungal species is small and determined by location, but not land use

Compared to bacteria and archaea, the fungal component of the soil microbiome (the mycobiome) was more variable across the different sampling sites. Indeed, out of 12,101 total OTUs, only 5 were present in all samples

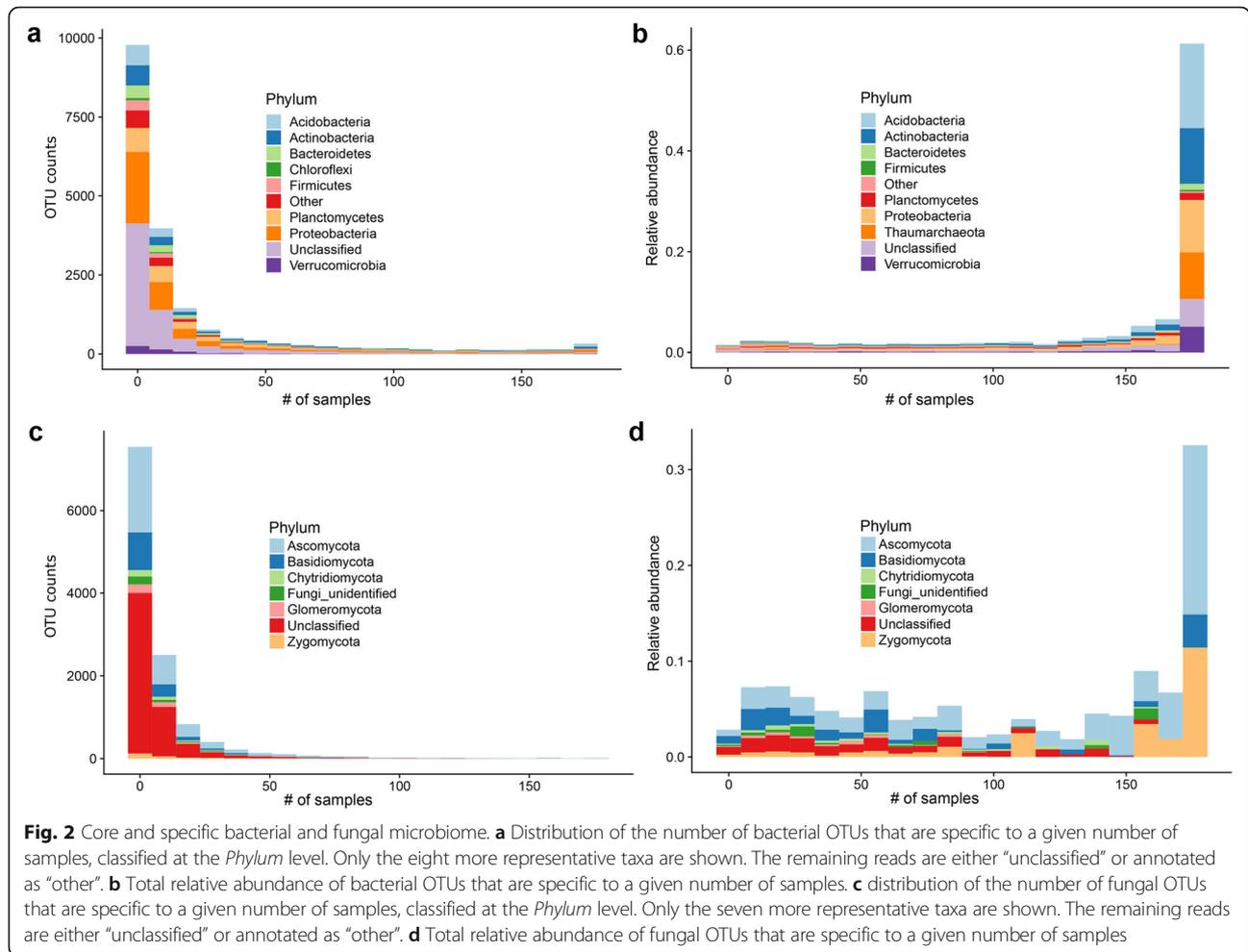


(Fig. 2c). Core OTUs accounted on average for  $15 \pm 5\%$  of each sample ( $35 \pm 8\%$  for OTUs present in more than 95% of the samples, Fig. 2d), while the specific OTUs (i.e., those present in less than 5% of the samples) represented an average of  $4.8 \pm 0.3\%$  of each mycobiome (Fig. 2d). From a taxonomic point of view, the core mycobiome was dominated by *Ascomycota*, *Zygomycota*, and a small fraction of *Basidiomycota*. Using the same subsampling strategy outlined above, the size of the core mycobiome as a function of the number of samples was again well described by a power-law (Additional file 6: Figure S6b) that in this case converged to a value close to zero ( $4.1 \pm 0.4$  OTUs). Differently to what found for bacteria, the core of conserved fungal species in the three cultivation types (17, 10, and 13 OTUs for V, P1, and P2 samples, respectively) was not significantly larger than expected by random sampling ( $p$  value 0.58). On the contrary, the different sites

had a site-specific core mycobiome (Additional file 17: Table S4) significantly larger than expected by random sampling ( $p$  value  $3.2 \times 10^{-4}$ ). These results suggest that geographical factors dominate the composition of the soil mycobiota that show a higher level of variability compared to bacteria, and that cultivation is not able to select a defined set of fungal species across different locations overcoming the differences due to geographical factors.

**Bacterial and fungal richness vary with site and cultivation, and are correlated in a site-specific manner**

The microbiome richness, or  $\alpha$ -diversity, varied widely across locations, sites, and cultivation type (Additional file 18: Table S5 and Additional file 19: Table S6) for both bacteria and fungi (Fig. 3a, b respectively). We first tested the differences across the locations (Additional file 18: Table S5) and sampling sites (Additional file 19: Table S6)



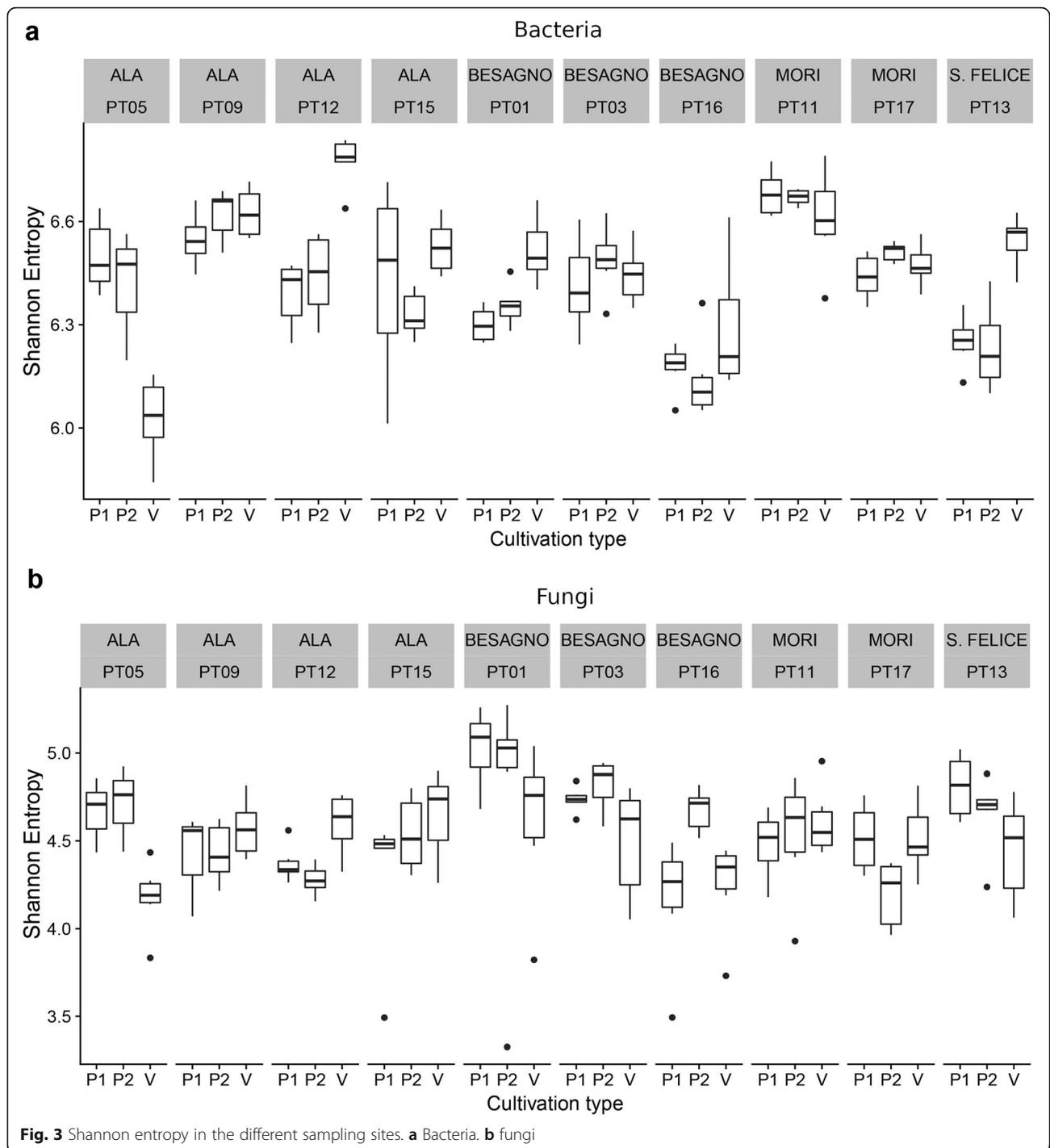
finding that in many cases the differences were significant. We then tested whether the microbiota richness was significantly different between the three cultures within the same sampling site. Despite the fact that samples from the vineyards had often a bacterial  $\alpha$ -diversity significantly different from surrounding permanent grassland, we could not highlight an unambiguous effect of cultivation across all sites (Additional file 20: Table S7). For bacteria, in no case the P1 and P2 samples had significantly different richness (exceeding the 0.05 significance level), while in three cases the V samples have a significantly higher, and in one lower richness than the corresponding P1 samples (Additional file 20: Table S7). For fungi, in two cases, the  $\alpha$ -diversity of the P2 samples were significantly different from the P1 samples, while in three cases, the V samples had a lower  $\alpha$ -diversity than the corresponding P1 samples, and in one case higher (Additional file 20: Table S7).

To highlight the possible effects of competition between bacteria and fungi, we tested by linear modeling whether there was a correlation between the bacterial and fungal  $\alpha$ -diversities in the same samples. We found that the

correlation across all samples was negligible (adjusted  $R^2 = -0.006$ ,  $p$  value = 0.86, Additional file 21: Table S8). However, correlating the diversity indexes within each site the correlation was significant (adjusted  $R^2 = 0.22$ ,  $p$  value =  $2.22 \times 10^{-6}$ , Additional file 22: Table S9). In most sites, the fungal  $\alpha$ -diversity was negatively correlated with the bacterial  $\alpha$ -diversity, with the exception of PT05 and PT12 both from the same area (Ala), where we found a positive correlation between bacterial and fungal  $\alpha$ -diversity ( $R^2 = 0.51$ , slope 0.91,  $p$  value = 0.0016, Additional file 23: Table S10 and  $R^2 = 0.48$ , slope 0.64,  $p$  value = 0.0008, Additional file 24: Table S11, respectively). For the other sites, using a linear mixed-effect model where the intercept was treated as random variable, we obtained an estimated value of  $-0.59$  ( $p$  value = 0.002, Additional file 25: Table S12) for the slope of the correlation between the fungal and bacterial  $\alpha$ -diversity.

#### Chemical characteristics of the soil partially explained the variability in richness

To identify soil features that have an impact on  $\alpha$ -diversity, we built a random forest model including the



texture (percentage of sand, silt, and clay) and chemical characteristics of the soil (absolute quantity of  $\text{CaCO}_3$ , Cu, Zn, Pb, Cd, soil organic matter, carbon-to-nitrogen ratio-C/N -, and pH). The model was able to account for 58.25% of the  $\alpha$ -diversity variability for bacteria and archaea. The more relevant characteristics were the absolute quantity of  $\text{CaCO}_3$ , the percentage of sand, and the

percentage of silt, followed by the absolute quantities of Zn and Cu, and by pH. Differently from what recently found in global surveys [7, 12, 45], the effect of pH was moderate, probably due to the relatively small range of pH values sampled in the present study. By linear modeling, we found that the bacterial microbiome richness was significantly positively correlated (Additional file 7:

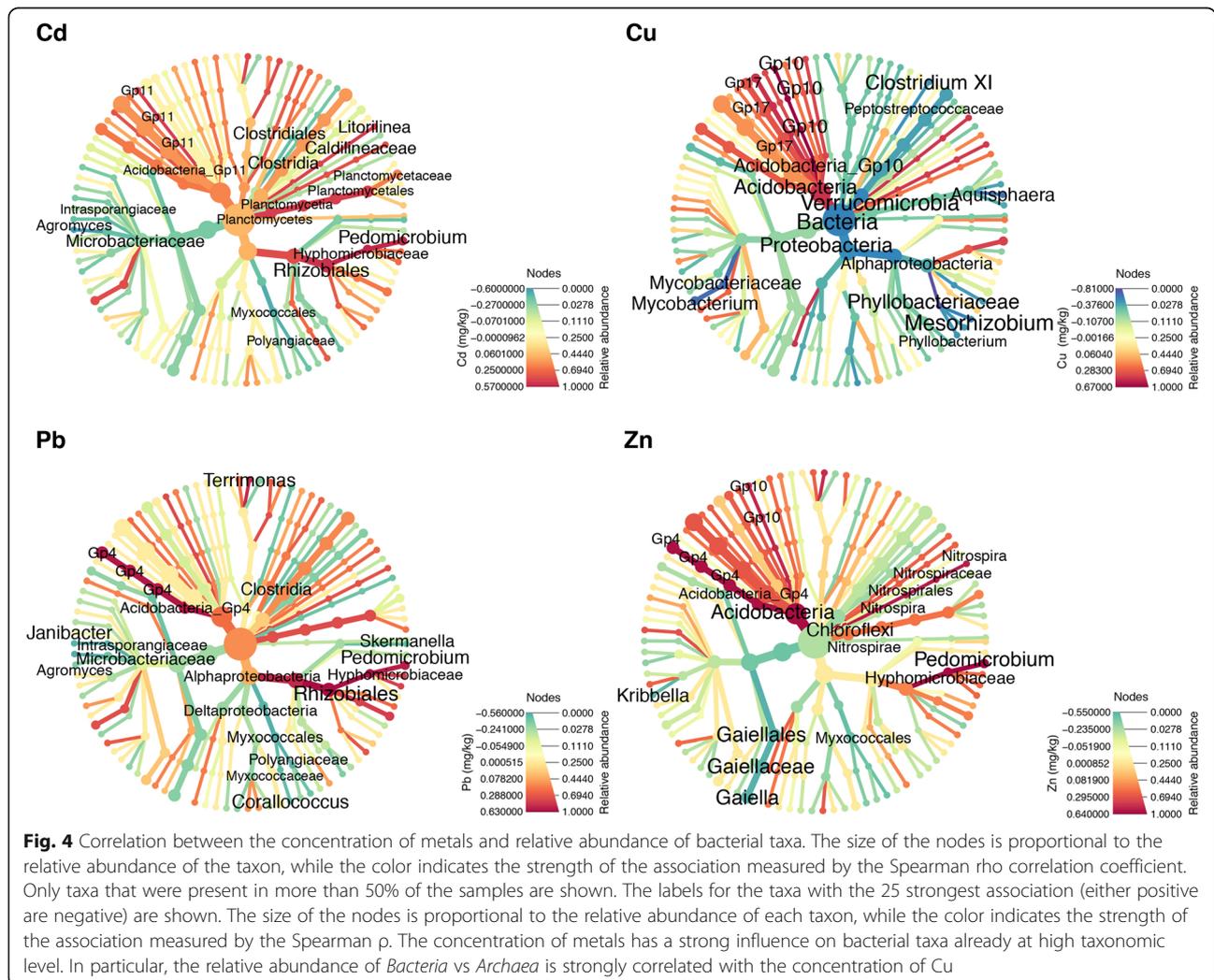
Figure S7 and Additional file 26: Table S13) with CaCO<sub>3</sub> and silt and negatively with sand, Cu, and Zn. The correlation with pH was not statistically significant.

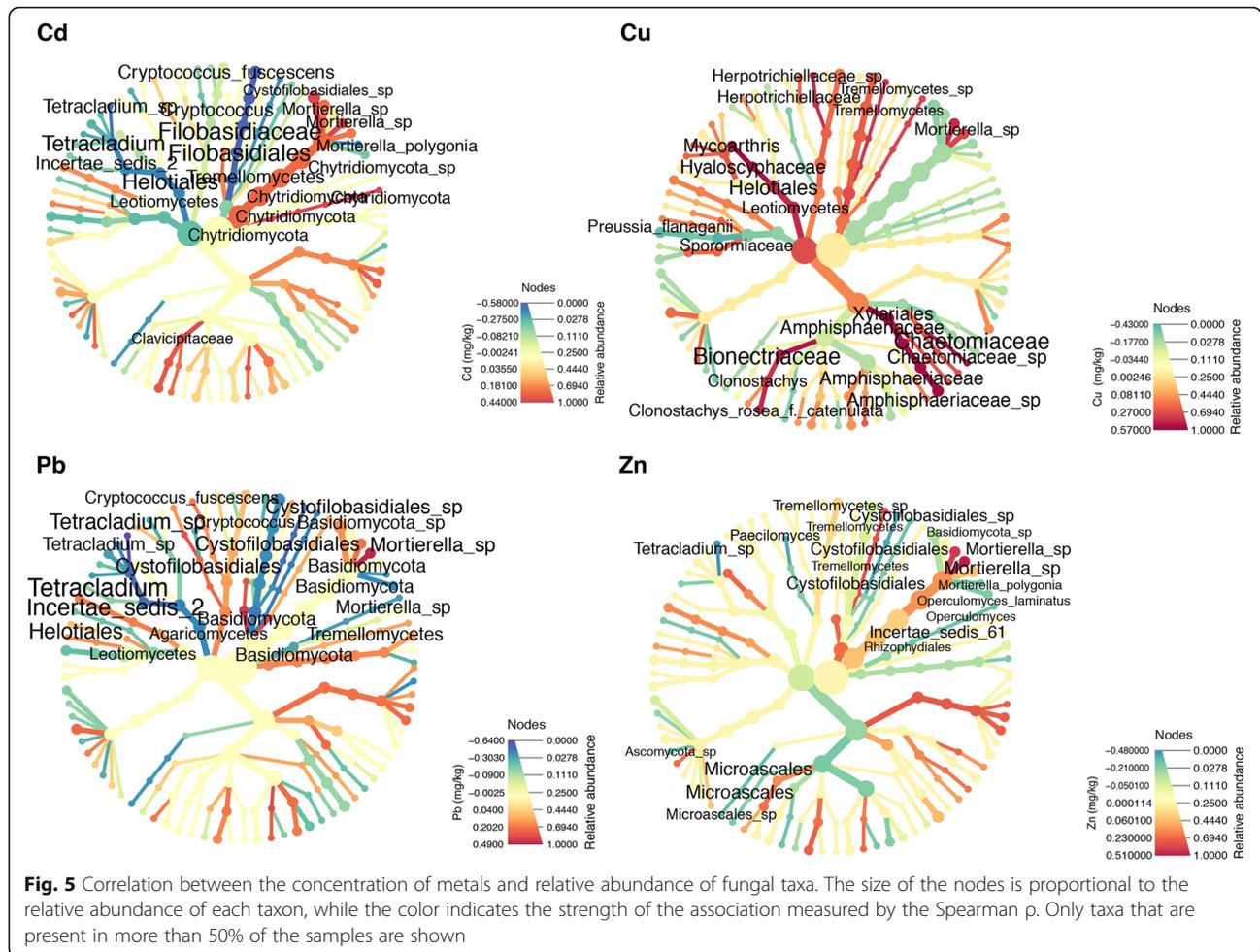
For fungi, a random forest model explained only 19.89% of the variability of the  $\alpha$ -diversity. The most relevant factor was the concentration of Cu, followed by the concentration of silt, Zn, total N, organic matter, and sand. A linear model showed that the concentration of Cu and Zn were negatively correlated to the  $\alpha$ -diversity of the fungal microbiota (Additional file 8: Figure S8 and Additional file 27: Table S14). The correlations with silt, total N, organic matter, and sand were not statistically significant.

**The characteristics of the soil had a large impact on the structure of the microbiota**

We performed a maximal information coefficient analysis [42] at all taxonomic ranks to identify taxa that were sensitive to specific characteristics of the soil (Additional file 28: Table S15 and Additional file 29: Table S16, respectively).

The characteristics of the soil correlated differently with different bacterial and fungal clades (Figs. 4, 5 and Additional file 9: Figure S9 and Additional file 10: Figure S10). The ratio between the relative abundances of bacteria and archaea was associated with all the measured quantities, with the exception of the C/N ratio and the concentration of Cd. In some cases (Additional file 11: Figure S11 and Additional file 28: Table S15), these associations correspond to non-zero values of Spearman's rho ( $\rho = 0.54$ , MICe = 0.38 for the concentration of Cu and  $\rho = -0.44$ , MICe = 0.37 for the relative abundance of sand, respectively), while in other cases, the Spearman's rho was close to 0 ( $\rho = -0.09$ , MICe = 0.22 for the relative abundance of clay). In general, the strongest associations (positive and negative, respectively) in 16S data are between the relative abundances of two bacterial families, namely, the family *Verrucomicrobiaceae* of the order *Verrucomicrobiales* [46] ( $\rho = 0.72$ , MICe = 0.51) that was positively correlated with the amount of silt, and the family *Phyllobacteriaceae* [47] of the order *Rhizobiales* ( $\rho = -0.81$ , MICe = 0.58), that was





negatively correlated with the concentration of Cu (Additional file 12: Figure S12 a and b, respectively).

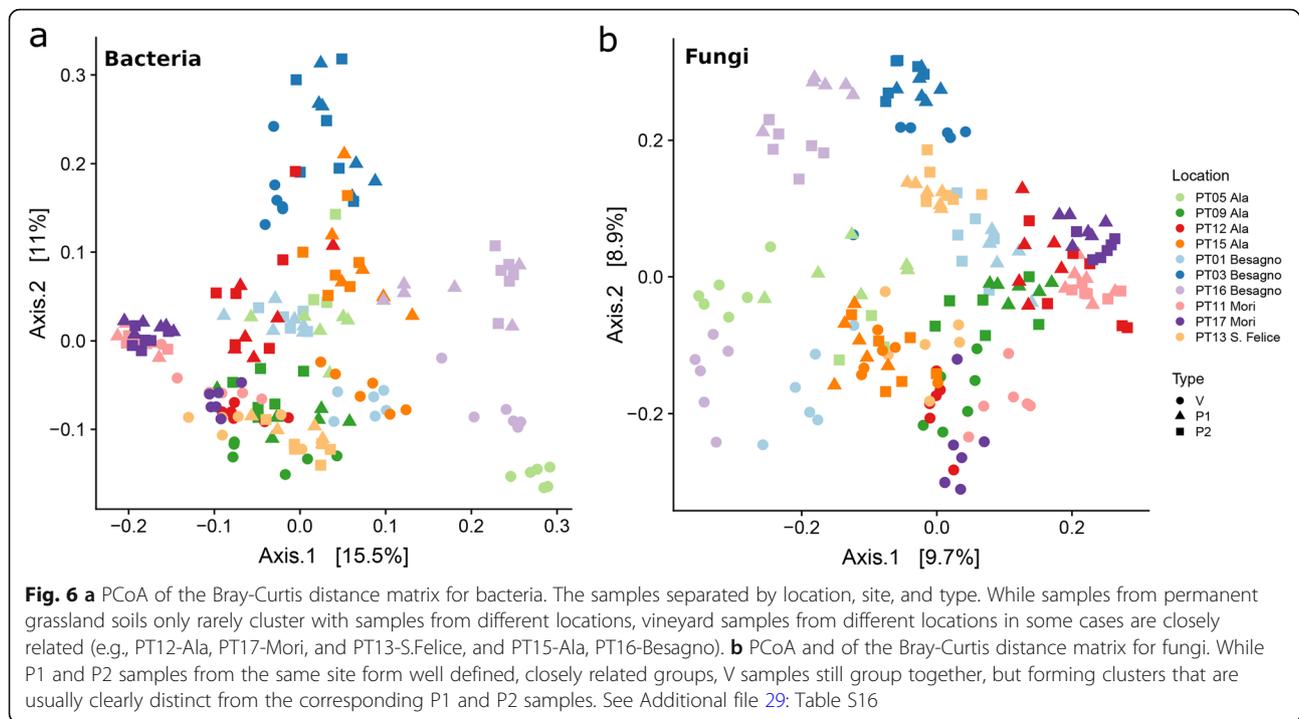
For fungi, several clades were strongly correlated to the characteristics of the soil (Fig. 5 and Additional file 10: Figure S10). The strongest associations were between one OTU from the family *Herpotrichiellaceae*, that was negatively correlated with the relative amount of silt ( $\rho = -0.80$ , MICe = 0.72) and positively with the relative amount of sand ( $\rho = 0.71$ , MICe = 0.55).

**The composition of the bacterial microbiome correlates with geography, while fungal communities are dominated by cultivation type**

To explore if differences in microbiome structure and composition correlate with sampling location and soil type, we computed the between-sample diversity (–diversity) using Bray-Curtis distance (Fig. 6). For bacteria, samples from the same location and type of soil generally clustered together (Fig. 6a). In addition, samples from both types of permanent grassland soil were closely related, with only one case in which a PERMANOVA test indicated a clear distinction (PT16, Additional file 29: Table S16) and two in which the

test was marginally significant ( $p$  value = 0.01, PT05 and PT17). Differently, in all locations, the vineyard soils were clearly distinct from the corresponding permanent grassland soils (Additional file 29: Table S16). However, in most cases, samples from the same location formed well-defined groups and the distances between permanent grassland and vineyard samples from the same site were smaller than the distances between samples from different sites (Additional file 13: Figure S13).

For fungi (Fig. 6b), samples from permanent grassland soils were closely related (with two exceptions, namely, PT01 and PT16, and one, PT17, for which the PERMANOVA test was marginally significant, i.e.,  $p = 0.01$ , Additional file 29: Table S16), and were distinct from corresponding vineyard samples. However, in contrast with what was found for bacteria, a large fraction of the samples from the vineyards formed in a large group (Fig. 6b), samples from the same site were in most cases similar to the distances between samples from different sites (Additional file 13: Figure S13). These results suggest that soil usage has a stronger effect on the fungal component of the soil microbiota than on bacteria.



## Discussion

Despite growing interest, the global and local variability of soil microbial communities and the biotic and abiotic factors that drive their differentiation are still poorly understood. Even on the local scale, soil can be composed by a large number of different habitats, that range from the immediate surroundings of plant roots (the rhizosphere) to bulk soil, where factors essential for survival of microbial communities (including, for instance, oxygen, water or nutrient availability) can vary substantially on the millimeter to centimeter scale [7, 10].

Soil microbial communities are of particular relevance in grape cultivation, since soil microbiota has been shown to serve as a reservoir of microorganisms colonizing grapes [4], contributing to shape regional wine characteristics [48].

To characterize the impact of cultivation and of other environmental factors on soil microbial communities in vineyards, we profiled via deep amplicon sequencing the bacterial and fungal components of the microbiome of soils sampled from vineyards and from associated permanent grasslands situated in close proximity in 10 sites geographically located in 4 different areas in the Adige Valley. Consistently with previous results obtained on both cultivated and permanent grassland soils [9, 49], a large fraction of the microbiome of each sample was constituted by sequences from yet uncharacterized taxa, thus supporting the notion that the soil environment is still severely underrepresented in sequence databases [50]. At the OTU level, we found a high degree of variability across samples. Most of the bacterial OTUs were present in a

small number of samples, and only a small number of core OTUs were shared by all samples. However, core OTUs were the dominant fraction of all the samples in terms of relative abundance, confirming earlier works [17] that have shown that approximately 500 bacterial phylotypes accounted for nearly half of the soil communities worldwide and supporting the recent finding that the bacterial diversity found on the local scale recapitulates what is found on the global scale [49]. The size of the core of shared bacterial OTUs increased significantly when we considered more homogenous samples, both in terms of geographical location or land use, showing that both dispersion limitations and agricultural practices influence the bacterial component of the bulk soil microbiota.

Surprisingly, we found that the variability of fungal communities followed patterns that were qualitatively different from what was found for bacteria. Using the ITS data from the same samples, we estimated that the core mycobiome was limited to less than 5 OTUs, accounting on average for approximately 15% of the sequences of each sample. Despite the relatively limited geographical range sampled in this study, these results are in striking agreement with the dominant role of a small number of taxa, in particular from the phylum *Ascomycota*, that has been highlighted by a global survey of soil samples [51]. In addition, we found that geographical location, but not land use, had an impact on determining the size of the core soil mycobiome, indicating the importance of spatial processes in structuring the biogeographic pattern of soil fungal communities [52] in

accordance with what found in grape associated bacterial and fungal communities [53], and confirming that dispersal limitations play a crucial role in determining the diversity across fungal communities even on a local geographical scale [8].

This difference between bacteria and fungi, the two dominant components of the soil microbiome [7], likely reflects the general characteristics of the different dispersal behaviors of these two classes of organisms. It has been previously suggested that the distribution of fungi exhibit strong biogeographic patterns that could be driven by dispersal limitations [13, 54], while bacteria are believed to have weak biogeographical patterns. Our results support the idea that while a core set of phylotypes that dominate soil bacterial microbiota from the local to the global scale [17], for fungi the set of core phylotypes is much more limited both in terms of number and of relative abundance, and is dominated by generalist taxa that can be disseminated by wind [51]. Comparison of the microbial diversity between vineyards and associated permanent grasslands suggests that the impact of cultivation on soil microbiota strongly depends on specific characteristics of the soil and possibly of the land management. The absence of a deterministic trend induced by different uses of similar soils is consistent with earlier studies that have found no general impact on the bacterial richness of long-term land-use change [15] and of cropping system [55]. However, in other cases, an increase in soil bacterial richness was found after the conversion of forest to agriculture [56], suggesting that the effect of cultivation strongly depends on the nature of the soil and cultivation type. For fungi, agricultural intensification has been shown to be the cause of reduced connectivity of interaction networks [57]. Despite the caution that should be used when interpreting co-occurrence networks inferred from microbiome data [58], this is an indication that farming destabilizing fungal soil communities, selecting some species over others [59] and possibly leading to a general instability of the community [60].

Interestingly, we identified a widespread (8 out of 10 sampling locations) inverse correlation between the diversity of the bacterial and fungal components of the soil microbiota. The correlations between fungal and bacterial richness in environmental samples have not been studied in detail, and likely depend both on the specific characteristics of the environment under study and on general mechanisms of interactions between these two groups of organisms. The number of fungal and bacterial OTUs was positively correlated in a study of the dust-associated microbiome on the continental scale [61]. Other studies have shown that bacterial and fungal diversity correlate in an opposite way to latitude on the global scale [12], and with other environmental variables in

soils in an alpine grassland ecosystem [62]. Recently, the correlation between antibiotic-resistance genes and the ratio between bacteria and fungi suggested a strong competition between the bacterial and fungal components of the microbiota through the production of antimicrobial substances from the latter [12]. The negative correlation between the fungal and bacterial diversity found here is consistent with a scenario where a diverse fungal community produces a range of antimicrobial molecules that in turn pose a strong selective pressure on bacteria [22]. Although definitive conclusions can be obtained only by correlating absolute abundances, and the fact that inferred microbial interactions can be biased by sample heterogeneity [63], our results highlight the role of fungal-bacterial interactions in determining the structure of the soil microbiota, suggesting that studies that aim at identifying the environmental factors that influence one of the two components should also take the other into account. The chemical and textural characteristics of soil partially determine the richness of the soil microbiota, and the link between the characteristics of the soil and the richness of the microbial communities is weaker for fungi than for bacteria. These results are consistent with earlier studies that have shown that the assembly of microbial communities is only partially determined by environmental conditions [54]. Individual taxa are influenced differently by physico-chemical characteristics of the soil, supporting the hypothesis that the microbiome structure can be manipulated [16].

Considering the between-samples variability, we found that the effect of land use on the structure and composition of the microbiome was especially strong for the fungal component, while for bacteria the geographical origin of the samples was the dominant factor. This result is in apparent contrast with the result that geography was the only relevant factor determining the size of the core of fungal species conserved across samples. However, the fungal component of the microbiota in samples from the vineyard was characterized by the presence of one OTU representing a large fraction of the sequences, distinguishing them from other samples without altering significantly the core size in terms of shared OTUs. This OTU was classified as a member of the *Amphisphaeriaceae*, a family of Ascomycota that includes both plant pathogens and endophytes [64, 65] and that has been shown to colonize grapevine wood [66] and survive to sterilization through hot water treatment [67]. This association with grapevine is probably the cause of the relatively high abundance of this fungus in vineyard samples, suggesting that the identity of the cultivated species, more than cultivation itself, has a strong influence on the fungal component of the soil microbiome.

## Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s40168-019-0758-7>.

**Additional file 1: Figure S1.** Schematic representation of sampling area. Each field was divided in 3 portions: the zone with grapevines (V) and two permanent grasslands, P1 and P2, respectively at 8 and 16 m from V; 6 replicates were made for each portion (black crosses). P1 and P2 represent portions of fields with a minimum human intervention, without tractors passage or chemical treatments, only mowing (twice per year). (JPG 231 kb)

**Additional file 2: Figure S2.** Geographical maps of sampling sites. The left panel reports all sampling sites compared to the whole Trentino-Alto Adige region, positions are then detailed in the right panel. The altitude respect to the sea level is represented as grayscale. "PT" prefix for sampling sites is omitted for sake of clarity. (JPG 735 kb)

**Additional file 3: Figure S3.** PCA of the chemical characteristics of the samples. (JPG 640 kb)

**Additional file 4: Figure S4.** a) Log<sub>2</sub> Fold Change of bacterial OTUs with relative abundances significantly different between vineyards and associated uncultivated fields. b) Relative abundances of the top ten bacterial OTUs whose relative abundances change most significantly between vineyards and associated permanent grassland. Only OTUs with more than 10 reads in 25% of the samples are shown. (JPG 2364 kb)

**Additional file 5: Figure S5.** a) Log<sub>2</sub> Fold Change of fungal OTUs with relative abundances significantly different between vineyards and associated permanent grassland. b) Relative abundances of the top ten fungal OTUs whose relative abundances change most significantly between vineyards and associated permanent grassland. Only OTUs with more than 10 reads in 25% of the samples are shown. (JPG 2076 kb)

**Additional file 6: Figure S6.** Size of the core bacterial a) and fungal b) microbiome as a function of the number of samples. The red line is a power law fit. (JPG 360 kb)

**Additional file 7: Figure S7.**  $\alpha$ -diversity of the bacterial microbiota as a function of the six most relevant features identified by random forest: CaCO<sub>3</sub> (A), Sand (B), Silt (C), Zinc (D), Copper (E) and pH (F). The blue line is a maximum likelihood linear fit, the shaded area is the 95% confidence interval of the linear model. Parameters of the fit are reported in Additional file 26: Table S13. (JPG 934 kb)

**Additional file 8: Figure S8.**  $\alpha$ -diversity of the fungal microbiota as a function of the six most relevant features identified by random forest: Copper (A), Silt (B), Zinc (C), Nitrogen (D), Organic matter (E), sand (F). The blue line is a maximum likelihood linear fit, the shaded area is the 95% confidence interval of the linear model. Parameters of the fit are reported in Additional file 27: Table S14. (JPG 922 kb)

**Additional file 9: Figure S9.** Correlation (Spearman rho) between the soil texture and chemical characteristics and the relative abundance of bacterial taxa. The size of the nodes is proportional to the relative abundance of the taxon, while the color indicates the strength of the association measured by the Spearman Rho correlation coefficient. Only taxa that are present in more than 50% of the samples are shown. (JPG 3518 kb)

**Additional file 10: Figure S10.** Correlation (Spearman rho) between the soil texture and chemical characteristics and the relative abundance of fungal taxa. The size of the nodes is proportional to the relative abundance of the taxon, while the color indicates the strength of the association measured by the Spearman Rho correlation coefficient. Only taxa that are present in more than 50% of the samples are shown. (JPG 3071 kb)

**Additional file 11: Figure S11.** Relative abundance of OTUs from the Kingdom *Bacteria* vs soil characteristics. Lines are *loess* smoothing of the data, shaded areas are the 95% confidence intervals around the smooth. (JPG 2712 kb)

**Additional file 12: Figure S12.** a) Relative abundance of the Family *Phyllobacteriaceae* as a function of the concentration of Cu. b) Relative abundance of the Family *Verrucomicrobiaceae* as a function of the concentration of Silt. The line is a *loess* smoothing of the data, and the shaded area is the 95% confidence interval. (JPG 726 kb)

**Additional file 13: Figure S13.** a) Boxplot of the Bray Curtis distances between the bacterial component of the microbiota of samples within the permanent grassland (P) and vineyard (V) soils and between vineyard and grassland soils (V vs P) for each site. b) same as a, across different sites. c) Same as a, for fungi. d) Same as b, for fungi. (JPG 47 kb)

**Additional file 14: Table S1.** Main features of sample fields (A and B). All samples were from vineyards of Adige valley in Trentino region and sampled the very same day (July the 12th of 2017). Legend. V = vineyard, P1 = grassland at 8 m from V, P2 = grassland at 16 m from V (see also Additional file 1: Figure S1). Loc. Name = Location Name, Trellis sys. = Trellis system, m asl = meters at sea level. (DOCX 15 kb)

**Additional file 15: Table S2.** Results of the Deseq analysis of bacterial taxa significantly differentially abundant in vineyard vs permanent grassland (P<sub>1</sub> and P<sub>2</sub> are considered together) samples for 16S data. (XLSX 53 kb)

**Additional file 16: Table S3.** Results of the Deseq analysis of fungal taxa significantly differentially abundant in vineyard vs permanent grassland (P<sub>1</sub> and P<sub>2</sub> are considered together) samples for ITS data. (XLSX 22 kb)

**Additional file 17: Table S4.** The site-specific core microbiome for Bacteria and Archaea, and Fungi. (DOCX 14 kb)

**Additional file 18: Table S5.** Pairwise comparison of location  $\alpha$ -diversity measured by the Shannon Entropy for bacteria A) and fungi B) using the Wilcoxon rank-sum test, FDR corrected. Highlighted in bold values below the significance threshold of 0.05. (DOCX 14 kb)

**Additional file 19: Table S6.** Pairwise comparison of site  $\alpha$ -diversity measured by the Shannon Entropy for bacteria a) and fungi b) using the Wilcoxon rank-sum test, FDR corrected. Comparisons are ordered by sampling area (PT13 is the only sample from S. Felice area). Highlighted in bold values below the significance threshold of 0.05 while light gray shade enlight comparisons inside the same area. (DOCX 17 kb)

**Additional file 20: Table S7.** Average differences of the  $\alpha$ -diversity of bacterial and fungal communities measured by the Shannon entropy between the Vineyards and P1 samples for each site. Statistically significant contrasts are highlighted in bold. While for bacteria in no case the P1 and P2 samples have significantly different richness, for fungi, in two cases the  $\alpha$ -diversity of the P2 samples are significantly different from the P1 samples (higher in PT16 and lower in PT17, *p*-values 0.00116 and 0.040, respectively). (DOCX 14 kb)

**Additional file 21: Table S8.** Linear model correlating the the bacterial and fungal  $\alpha$ -diversities in all sites. (DOCX 14 kb)

**Additional file 22: Table S9.** Linear model correlating the the bacterial and fungal  $\alpha$ -diversities, stratified by site. (DOCX 14 kb)

**Additional file 23: Table S10.** Linear model correlating the the bacterial and fungal  $\alpha$ -diversities for PT05. (DOCX 13 kb)

**Additional file 24: Table S11.** Linear model correlating the the bacterial and fungal  $\alpha$ -diversities for PT12. (DOCX 13 kb)

**Additional file 25: Table S12.** Linear mixed effect model correlating the the bacterial and fungal  $\alpha$ -diversities for all sites except PT05 and PT12. Intercept is a treated as a random effect, slope as a fixed effect. (DOCX 13 kb)

**Additional file 26: Table S13.** Parameters of the linear models in Additional file 6: Figure S6 modeling the richness of bacterial microbiota (Shannon entropy) against the chemical characteristics of the soil. (DOCX 14 kb)

**Additional file 27: Table S14.** Parameters of the linear models in Additional file 7: Figure S7 modeling the richness of fungal microbiota (Shannon entropy) against the chemical characteristics of the soil. (DOCX 14 kb)

**Additional file 28: Table S15.** Taxa significantly correlated with chemical and physical characteristics of the soil from 16S data and ITS data. (XLSX 654 kb)

**Additional file 29: Table S16.** PERMANOVA results for the distinction between P1 and P2 (left) and P1 + P2 and V (right) for bacteria and fungi. In the P1 + P2 vs V comparison the P1 and P2 samples were considered together. Significant *p*-values (*p* < 0.01) are marked in bold. (DOCX 14 kb)

**Acknowledgements**

Not applicable.

**Authors' contributions**

EC defined the experimental plan, carried out the sample collection, data analysis, and contributed to the manuscript. AC helped to define the experimental plan, carried out the sample collection, and contributed to the manuscript. RZ carried out the sample collection, the laboratory work, and contribute to the manuscript. DB, MP, and SL carried out the laboratory work. DA made the data analysis. EM defined the experimental plan and carried out the sample collection. CD defined the experimental plan, made the data analysis, and wrote the manuscript. All authors edited the manuscript and approved the final draft.

**Funding**

This work was funded by the "Provincia Autonoma di Trento" in the framework of "Accordi di Programma" (ADP P16110511) and with the contribution of "Consorzio Vini del Trentino" (<https://www.vinideltrentino.com/>).

**Availability of data and materials**

Raw sequencing data are available at the European Nucleotide Archive (<https://www.ebi.ac.uk/ena>) under the study id PRJEB31356.

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Faculty of Science and Technology, Free University of Bozen, Pza Università 1, 39100 Bolzano, Italy. <sup>2</sup>Unit of Computational Biology, Research and Innovation Centre, Fondazione Edmund Mach, Via Mach 1, 38010 San Michele all'Adige, Italy. <sup>3</sup>Technology Transfer Center, Fondazione Edmund Mach, Via Mach 1, 38010 San Michele all'Adige, Italy.

Received: 4 March 2019 Accepted: 9 October 2019

Published online: 08 November 2019

**References**

- Woodcroft BJ, Singleton CM, Boyd JA, Evans PN, Emerson JB, Zayed AAF, et al. Genome-centric view of carbon processing in thawing permafrost. *Nature*. 2018;560:49–54. <https://doi.org/10.1038/s41586-018-0338-1>.
- Crits-Christoph A, Diamond S, Butterfield CN, Thomas BC, Banfield JF. Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. *Nature*. 2018;558:440–4. <https://doi.org/10.1038/s41586-018-0207-y>.
- Wall DH, Nielsen UN, Six J. Soil biodiversity and human health. *Nature*. 2015;528:69–76. <https://doi.org/10.1038/nature15744>.
- Zarraonaindia I, Owens SM, Weisenhorn P, West K, Hampton-Marcell J, Lax S, et al. The soil microbiome influences grapevine-associated microbiota. *MBio*. 2015;6. <https://doi.org/10.1128/mBio.02527-14>.
- Laforest-Lapointe I, Paquette A, Messier C, Kembel SW. Leaf bacterial diversity mediates plant diversity and ecosystem function relationships. *Nature*. 2017;546:145–7. <https://doi.org/10.1038/nature22399>.
- Hacquard S, Spaepen S, Garrido-Oter R, Schulze-Lefert P. Interplay between innate immunity and the plant microbiota. *Annu Rev Phytopathol*. 2017;55:565–89. <https://doi.org/10.1146/annurev-phyto-080516-035623>.
- Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol*. 2017;15:579–90. <https://doi.org/10.1038/nrmicro.2017.87>.
- Peay KG, Kennedy PG, Talbot JM. Dimensions of biodiversity in the earth microbiome. *Nat Rev Microbiol*. 2016;14:434–47. <https://doi.org/10.1038/nrmicro.2016.59>.
- Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*. 2017;551:457–63. <https://doi.org/10.1038/nature24621>.
- O'Brien SL, Gibbons SM, Owens SM, Hampton-Marcell J, Johnston ER, Jastrow JD, et al. Spatial scale drives patterns in soil bacterial diversity. *Environ Microbiol*. 2016;18:2039–51. <https://doi.org/10.1111/1462-2920.13231>.
- Meyer KM, Memiaghe H, Korte L, Kenfack D, Alonso A, Bohannan BJM. Why do microbes exhibit weak biogeographic patterns? *ISME J*. 2018;12:1404–13. <https://doi.org/10.1038/s41396-018-0103-3>.
- Bahram M, Hildebrand F, Forslund SK, Anderson JL, Soudzilovskaia NA, Bodegom PM, et al. Structure and function of the global topsoil microbiome. *Nature*. 2018;560:233–7. <https://doi.org/10.1038/s41586-018-0386-6>.
- Tedersoo L, Bahram M, Pölme S, Kõljalg U, Yorou NS, Wijesundera R, et al. Fungal biogeography. Global diversity and geography of soil fungi. *Science*. 2014;346:1256688. <https://doi.org/10.1126/science.1256688>.
- Maestre FT, Delgado-Baquerizo M, Jeffries TC, Eldridge DJ, Ochoa V, Gozalo B, et al. Increasing aridity reduces soil microbial diversity and abundance in global drylands. *Proc Natl Acad Sci U S A*. 2015;112:15684–9. <https://doi.org/10.1073/pnas.1516684112>.
- Goss-Souza D, Mendes LW, Borges CD, Baretta D, Tsai SM, Rodrigues JLM. Soil microbial community dynamics and assembly under long-term land use change. *FEMS Microbiol Ecol*. 2017;93. <https://doi.org/10.1093/femsec/fix109>.
- Leff JW, Jones SE, Prober SM, Barberán A, Borer ET, Firn JL, et al. Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proc Natl Acad Sci U S A*. 2015;112:10967–72. <https://doi.org/10.1073/pnas.1508382112>.
- Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett RD, et al. A global atlas of the dominant bacteria found in soil. *Science*. 2018;359:320–5. <https://doi.org/10.1126/science.aap9516>.
- Lynch MDJ, Neufeld JD. Ecology and exploration of the rare biosphere. *Nat Rev Microbiol*. 2015;13:217–29. <https://doi.org/10.1038/nrmicro3400>.
- Chou M-Y, Vanden Heuvel J, Bell TH, Panke-Buisse K, Kao-Kniffin J. Vineyard under-vine floor management alters soil microbial composition, while the fruit microbiome shows no corresponding shifts. *Sci Rep*. 2018;8:11039. <https://doi.org/10.1038/s41598-018-29346-1>.
- Knight S, Klaere S, Fedrizzi B, Goddard MR. Regional microbial signatures positively correlate with differential wine phenotypes: evidence for a microbial aspect to terroir. *Sci Rep*. 2015;5:14233. <https://doi.org/10.1038/srep14233>.
- Gilbert JA, van der Lelie D, Zarraonaindia I. Microbial terroir for wine grapes. *Proc Natl Acad Sci U S A*. 2014;111:5–6. <https://doi.org/10.1073/pnas.1320471110>.
- Shaw LP, Bassam H, Barnes CP, Walker AS, Klein N, Balloux F. Modelling microbiome recovery after antibiotics using a stability landscape framework. *ISME J*. 2019;13:1845–56. <https://doi.org/10.1038/s41396-019-0392-1>.
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A*. 2011;108(Suppl 1):4516–22. <https://doi.org/10.1073/pnas.1000080107>.
- Apprill A, McNally S, Parsons R, Weber L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat Microb Ecol*. 2015;75:129–37. <https://doi.org/10.3354/ame01753>.
- Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol*. 2016;18:1403–14. <https://doi.org/10.1111/1462-2920.13023>.
- Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, et al. Improved bacterial 16S rRNA gene (V4 and V4–5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems*. 2016;1. <https://doi.org/10.1128/mSystems.00009-15>.
- Gardes M, Bruns TD. ITS primers with enhanced specificity for basidiomycetes—application to the identification of mycorrhizae and rusts. *Mol Ecol*. 1993;2:113–8. <https://doi.org/10.1111/j.1365-294X.1993.tb00005.x>.
- White TJ, Bruns T, Lee S, Taylor J. Amplification and direct sequencing of fungal ribosomal rna genes for phylogenetics. Editor(s): Innis MA, Gelfand DH, Sninsky JJ, White TJ. In: PCR protocols. Elsevier: Academic Press; 1990. p. 315–322. <https://doi.org/10.1016/B978-0-12-372180-8.50042-1>.
- Albanese D, Fontana P, De Filippo C, Cavalieri D, Donati C. MICCA: a complete and accurate software for taxonomic profiling of metagenomic data. *Sci Rep*. 2015;5:9743. <https://doi.org/10.1038/srep09743>.
- Edgar RC, Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*. 2015;31:3476–82. <https://doi.org/10.1093/bioinformatics/btv401>.
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench

- compatible with ARB. *Appl Environ Microbiol.* 2006;72:5069–72. <https://doi.org/10.1128/AEM.03006-05>.
32. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ.* 2016;4:e2584. <https://doi.org/10.7717/peerj.2584>.
  33. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol.* 2007;73:5261–7. <https://doi.org/10.1128/AEM.00062-07>.
  34. DeSantis TZ, Hugenholtz P, Keller K, Brodie EL, Larsen N, Piceno YM, et al. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Res.* 2006;34(Web Server issue):W394–9. <https://doi.org/10.1093/nar/gkl244>.
  35. Price MN, Dehal PS, Arkin AP. FastTree 2 — approximately maximum-likelihood trees for large alignments. *PLoS One.* 2010;5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
  36. Kõljalg U, Larsson K-H, Abarenkov K, Nilsson RH, Alexander IJ, Eberhardt U, et al. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. *New Phytol.* 2005;166:1063–8. <https://doi.org/10.1111/j.1469-8137.2005.01376.x>.
  37. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE.* 2013;8:e61217. <https://doi.org/10.1371/journal.pone.0061217>.
  38. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J.* 1948;27:379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
  39. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
  40. Hawinkel S, Mattiello F, Bijmans L, Thas O. A broken promise: microbiome differential abundance methods do not control the false discovery rate. *Brief Bioinform.* 2019;20:210–21. <https://doi.org/10.1093/bib/bbx104>.
  41. Liaw A, Wiener M. Classification and regression by random forest. *R News.* 2002;2:18–22.
  42. Albanese D, Riccadonna S, Donati C, Franceschi P. A practical tool for maximal information coefficient analysis. *Gigascience.* 2018;7:1–8. <https://doi.org/10.1093/gigascience/giy032>.
  43. Reshef YA, Reshef DN, Finucane HK, Sabeti PC, Mitzenmacher M. Measuring dependence powerfully and equitably. *J Mach Learn Res.* 2016;17:1–63.
  44. Foster ZSL, Shapton TJ, Grünwald NJ. Metacoder: an R package for visualization and manipulation of community taxonomic diversity data. *PLoS Comput Biol.* 2017;13:e1005404. <https://doi.org/10.1371/journal.pcbi.1005404>.
  45. Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG, et al. Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J.* 2010;4:1340–51. <https://doi.org/10.1038/ismej.2010.58>.
  46. Navarrete AA, Soares T, Rossetto R, van Veen JA, Tsai SM, Kuramae EE. Verrucomicrobial community structure and abundance as indicators for changes in chemical factors linked to soil fertility. *Antonie Van Leeuwenhoek.* 2015;108:741–52. <https://doi.org/10.1007/s10482-015-0530-3>.
  47. Gopalakrishnan S, Sathya A, Vijayabharathi R, Varshney RK, Gowda CLL, Krishnamurthy L. Plant growth promoting rhizobia: challenges and opportunities. *3 Biotech.* 2015;5:355–77. <https://doi.org/10.1007/s13205-014-0241-x>.
  48. Bokulich NA, Collins TS, Masarweh C, Allen G, Heymann H, Ebeler SE, et al. Associations among wine grape microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics. *MBio.* 2016;7. <https://doi.org/10.1128/mBio.00631-16>.
  49. Ramirez KS, Leff JW, Barberán A, Bates ST, Betley J, Crowther TW, et al. Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. *Proc Biol Sci.* 2014;281. <https://doi.org/10.1098/rspb.2014.1988>.
  50. Schloss PD, Girard RA, Martin T, Edwards J, Thrash JC. Status of the archaeal and bacterial census: an update. *MBio.* 2016;7. <https://doi.org/10.1128/mBio.00201-16>.
  51. Egdí E, Delgado-Baquerizo M, Plett JM, Wang J, Eldridge DJ, Bardgett RD, et al. A few Ascomycota taxa dominate soil fungal communities worldwide. *Nat Commun.* 2019;10:2369. <https://doi.org/10.1038/s41467-019-10373-z>.
  52. Peay KG, Bidartondo MI, Elizabeth AA. Not every fungus is everywhere: scaling to the biogeography of fungal-plant interactions across roots, shoots and ecosystems. *New Phytol.* 2010;185:878–82. <https://doi.org/10.1111/j.1469-8137.2009.03158.x>.
  53. Miura T, Sánchez R, Castañeda LE, Godoy K, Barbosa O. Is microbial terroir related to geographic distance between vineyards? *Environ Microbiol Rep.* 2017;9:742–9. <https://doi.org/10.1111/1758-2229.12589>.
  54. Schmidt SK, Nemergut DR, Darcy JL, Lynch R. Do bacterial and fungal communities assemble differently during primary succession? *Mol Ecol.* 2014;23:254–8. <https://doi.org/10.1111/mec.12589>.
  55. Hartman K, van der Heijden MGA, Wittwer RA, Banerjee S, Walser J-C, Schlaeppi K. Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. *Microbiome.* 2018;6:14. <https://doi.org/10.1186/s40168-017-0389-9>.
  56. Mendes LW, de Lima Bossi MJ, Kuramae EE, Tsai SM. Land-use system shapes soil bacterial communities in southeastern Amazon region. *Appl Soil Ecol.* 2015;95:151–60. <https://doi.org/10.1016/j.apsoil.2015.06.005>.
  57. Banerjee S, Walder F, Büchi L, Meyer M, Held AY, Gättinger A, et al. Agricultural intensification reduces microbial network complexity and the abundance of keystone taxa in roots. *ISME J.* 2019;13:1722–36. <https://doi.org/10.1038/s41396-019-0383-2>.
  58. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome datasets are compositional: and this is not optional. *Front Microbiol.* 2017;8:2224. <https://doi.org/10.3389/fmicb.2017.02224>.
  59. Verbruggen E, Rillig MC, Wehner J, Hegglin D, Wittwer R, van der Heijden MGA. Sebaciniales, but not total root associated fungal communities, are affected by land-use intensity. *New Phytol.* 2014;203:1036–40. <https://doi.org/10.1111/nph.12884>.
  60. Barabási A-L. Scale-free networks: a decade and beyond. *Science.* 2009;325:412–3. <https://doi.org/10.1126/science.1173299>.
  61. Barberán A, Ladau J, Leff JW, Pollard KS, Menninger HL, Dunn RR, et al. Continental-scale distributions of dust-associated bacteria and fungi. *Proc Natl Acad Sci U S A.* 2015;112:5756–61. <https://doi.org/10.1073/pnas.1420815112>.
  62. Zhang Y, Dong S, Gao Q, Liu S, Ganjurjav H, Wang X, et al. Soil bacterial and fungal diversity differently correlated with soil biochemistry in alpine grassland ecosystems in response to environmental changes. *Sci Rep.* 2017;7:43077. <https://doi.org/10.1038/srep43077>.
  63. Armitage DW, Jones SE. How sample heterogeneity can obscure the signal of microbial interactions. *ISME J.* 2019. <https://doi.org/10.1038/s41396-019-0463-3>.
  64. Jeewon R, Liew ECY, Hyde KD. Molecular systematics of the Amphisphaeriaceae based on cladistic analyses of partial LSU rDNA gene sequences. *Mycol Res.* 2003;107:1392–402. <https://doi.org/10.1017/S095375620300875X>.
  65. Hofstetter V, Buyck B, Croll D, Viret O, Couloux A, Gindro K. What if esca disease of grapevine were not a fungal disease? *Fungal Divers.* 2012;54:51–67. <https://doi.org/10.1007/s13225-012-0171-z>.
  66. Jayawardena RS, Purahong W, Zhang W, Wubet T, Li X, Liu M, et al. Biodiversity of fungi on *Vitis vinifera* L. revealed by traditional and high-resolution culture-independent approaches. *Fungal Divers.* 2018;90:1–84. <https://doi.org/10.1007/s13225-018-0398-4>.
  67. Eichmeier A, Pečenka J, Peňázová E, Baránek M, Català-García S, León M, et al. High-throughput amplicon sequencing-based analysis of active fungal communities inhabiting grapevine after hot-water treatments reveals unexpectedly high fungal diversity. *Fungal Ecol.* 2018;36:26–38. <https://doi.org/10.1016/j.funeco.2018.07.011>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

