

Article

# Evaluation of Soil Biodiversity in Alpine Habitats through eDNA Metabarcoding and Relationships with Environmental Features

Noemi Rota <sup>1,\*</sup>, Claudia Canedoli <sup>1</sup>, Chiara Ferrè <sup>1</sup>, Gentile Francesco Ficetola <sup>2,3</sup>, Alessia Guerrieri <sup>2</sup> and Emilio Padoa-Schioppa <sup>1</sup>

<sup>1</sup> Department of Earth and Environmental Sciences, University of Milano-Bicocca, 1-20126 Milan, Italy; claudia.canedoli@unimib.it (C.C.); chiara.ferre@unimib.it (C.F.); emilio.padoaschioppa@unimib.it (E.P.-S.)

<sup>2</sup> Dipartimento di Scienze e Politiche Ambientali, Università degli Studi di Milano, via Celoria 26, I-20133 Milano, Italy; francesco.ficetola@gmail.com (G.F.F.); alessia.guerrieri@unimi.it (A.G.)

<sup>3</sup> LECA, Laboratoire d'Ecologie Alpine, Université Grenoble Alpes, CNRS, Université Savoie Mont Blanc, F-38000 Grenoble, France

\* Correspondence: rota.noemi@gmail.com

Received: 21 May 2020; Accepted: 4 July 2020; Published: 7 July 2020



**Abstract:** Soil biodiversity is fundamental for ecosystems, ensuring many ecosystem functions, such as nutrient cycling, organic matter decomposition, soil formation, and organic carbon pool increase. Due to these roles, there is a need to study and completely understand how soil biodiversity is composed through different habitats. The aim of this study was to describe the edaphic soil community of the alpine environments belonging to the Gran Paradiso National Park, thus detecting if there are any correlation with environmental features. We studied soil fauna through environmental DNA metabarcoding. From eDNA metabarcoding, 18 families of arthropods were successfully detected, and their abundance expressed in terms of the relative frequency of sequences. Soil faunal communities of mixed coniferous forests were characterized by Isotomidae, Entomobryidae, Hypogastruridae, and Onychiuridae; while mixed deciduous forests were composed mostly by Isotomidae, Cicadidae, Culicidae, and Neelidae. Calcicolous and acidic grasslands also presented families that were not detected in forest habitats, in particular Scarabaeidae, Curculionidae, Brachyceridae, and had in general a more differentiated soil community. Results of the Canonical Component Analysis revealed that the main environmental features affecting soil community for forests were related to vegetation (mixed deciduous forests, tree basal area, tree biomass, Shannon index), soil (organic layers and organic carbon stock), and site (altitude); while for prairies, soil pH and slope were also significant in explaining soil community composition. This study provided a description of the soil fauna of alpine habitats and resulted in a description of community composition per habitat and the relation with the characteristic of vegetation, soil, and topographic features of the study area. Further studies are needed to clarify ecological roles and needs of these families and their role in ecosystem functioning.

**Keywords:** Gran Paradiso National Park; Alps; pedofauna; soil community; forests; prairies

## 1. Introduction

### 1.1. Soil Biodiversity and Ecosystem Functions

Soil hosts a huge amount of biodiversity, which performs an important role in the ecosystem functions. The current understanding of soil biodiversity is still quite limited because soil is highly heterogeneous and the organisms that live in soil are very small, difficult to isolate, and have a great range of habitats [1].

Soil invertebrates are the most abundant and differentiated animal group. They occupy many niches and their populations, phenology, behavior, dispersal, and genetics are highly sensitive to climate change [2]. Soil invertebrates ensure numerous important functions that benefit the ecosystems to which they belong. They are essential to the detritus food chain [3] and on the whole, the soil food web, and contribute to important ecological functions, such as organic matter decomposition, nutrient recycling, and water retention [4–7].

Soil fauna might influence Soil Organic Carbon (SOC) storage in different ways. For example, it might affect the soil structure or it might have a role in SOC accumulation and stabilization [8,9]. Soil macrofauna (e.g., earthworms) breaks down and transforms particulate organic matter, and can translocate SOC to greater soil depths, where it will have longer residence times [9,10]. However, limited information is available about the relationship between soil fauna and SOC storage. The edaphic faunal contribution to SOC dynamics depends on a great number of factors and might be significantly influenced by site-specific properties [11].

The composition of the arthropod communities present in soil varies widely from habitat to habitat (e.g., from grasslands to forests). Multiple environmental factors contribute to this variation, including differences in the organic matter present in soil [12].

Since soil invertebrates are inherently important, both because of their contributions to biodiversity and the ecological benefits they offer, it is essential to develop a better understanding of soil biodiversity, especially at the species and population levels. Even though databases with information on this topic are growing in number (Barcode of Life Data Systems–BOLD), information about these taxa is still lacking, and they urgently need to be studied.

Once soil biodiversity is properly understood, studies of the ecosystem functions should be performed. Such studies will improve our understanding of soil-based ecosystems functions and give us a first evaluation of ecosystem services correlated. In regards to concerns about the environment, it is important to note that soil invertebrate communities are deeply affected by human activities [13,14]. Studying such communities will lead to a better valuation of the conservation status of these taxa and will help clarify whether conservation actions are needed or not.

### *1.2. A New Approach to Pedofauna Description and Its Relationship with Environmental Features and Ecosystem Functions*

In this study, we used environmental DNA (eDNA) metabarcoding to characterize the composition of arthropod, insect, and annelid communities in alpine soils.

Environmental DNA is the mixture of DNA left in the environment by organisms, and it can be extracted from an environmental sample (e.g., soil, sediment, water, etc.) [15].

From a single environmental sample [16], eDNA metabarcoding allows the extraction and amplification of multiple taxonomic groups. Therefore, eDNA metabarcoding is useful for several tasks, including biodiversity description, monitoring the effects of climate change on ecosystems, and microbial analysis [17]. There are many additional advantages to eDNA metabarcoding. It can be implemented quickly [18,19]. It is non-invasive and eliminates the need to sample whole organisms [20]. It also allows to develop an exhaustive description of soil communities from different habitats, much faster than time-consuming traditional approaches, such as expert-based taxonomic identification. Furthermore, research done with eDNA barcoding is not problematized by the types of human errors linked to expert based techniques [18]. Nevertheless, an effective application of this technique requires a robust experimental design and the careful implementation of quality-checks across every step of the analyses [21–23]. Quality checks help avoid biased inferences, which are the false identification of organisms that are not actually present (false positives) or the non-identification of organisms that are actually present (false negatives) [23].

A growing number of studies are using eDNA barcoding to study soil biodiversity and improve our understanding of below-ground diversity.

Alpine areas are sensitive to natural and anthropogenic environmental changes [24]. The latter can be related to landscape management, unsustainable tourism, and climate change [25]. Due to their inherent sensitivity, alpine ecosystems need to be studied. Studies such as this one provide policy-makers with the tools they need to manage and protect natural areas, evaluate their conservation status, and gain a better understanding of their ecological functions.

The aim of the study was to describe pedofauna communities through eDNA metabarcoding in the Gran Paradiso National Park (Italy). This description was done at two depth layers of soil, in order to describe vertical distribution of the edaphic soil biodiversity. The studied layers ranged from the surface to below-ground and were located in the six most extended habitats in the study area. After gathering the data from the field, the second aim was to relate differences in soil community composition to environmental features, such as vegetation type, soil characteristics, and study area features. This investigation also led to the development of considerations regarding the relationship between biodiversity and environmental features.

## 2. Materials and Methods

### 2.1. Study Area

The study took place in the Italian Alps at the Gran Paradiso National Park, the first protected area established in Italy in 1922. The Park is situated in Northwestern Italy. It occupies the regions of the Aosta Valley and Piedmont. It extends for nearly 600 km<sup>2</sup>, and its altitudes range from 800 m to 4000 m. Gran Paradiso National Park has five valleys: Val Soana, Valle Orco, Val di Rhemes, Valsavaranche, and the Val di Cogne. The geomorphology of the Park derives from glacial processes: it is due to the expansion of glaciers during Quaternary glaciations.

### 2.2. Sampling

Soil and vegetation samples were collected from June to August during the summers of 2018 and 2019, in four of the five valleys belonging to the park. A total of 122 plots (each 30 × 30 m; Figure S1) were established, with a stratified random sampling design distributed across the most extended habitats of the forests (70 plots) and the grasslands (52 plots) of the park (Table S1), in order to yield a more representative sampling of the study area. Eighty of the 122 plots were investigated with eDNA techniques (Table S2). The process of plot selection was highly based on the habitat map of the park, which provide the pre-existent habitat classifications. The various habitats of the park were identified and quantified in terms of their area, and only the most spatially extensive forests and grasslands were chosen for study (Table 1). The primary types of forest in the park are mixed deciduous, mixed coniferous, spruce (composed mostly by *Picea abies* L.), and larch (composed of *Larix decidua* Mill.). Meanwhile, the major categories of grasslands are acidic grassland and calcicolous grassland. The park also contains rocky environments, glaciers, shrub environments, and aquatic environments. These areas were not included in this study because of their limited geographical extension or the scarce presence of soil.

**Table 1.** Extension of habitat typologies of forests and grasslands in the Park.

	Habitat Type	Area (ha)	Used in This Study
Forests	Larch ( <i>Larix decidua</i> )	6890	*
	Mixed deciduous	2455	*
	Mixed coniferous	913	*
	Spruce ( <i>Picea abies</i> )	394	*
	Sparse coniferous	143	
	Reforestations	130	

Table 1. Cont.

	Habitat Type	Area (ha)	Used in This Study
Forests	Isolated trees or in small groups	120	
	Durmast ( <i>Quercus petraea</i> Matt.) <i>Que</i>	89	
	Fir ( <i>Abies</i> Mill.)	63	
	Mixed coniferous and deciduous	62	
	Pine ( <i>Pinus Sylvestris</i> L.)	41	
	Pine ( <i>Pinus mugo</i> Turr.)	21	
	Mixed hygrophile deciduous	2	
Grasslands	Cores of standing dead plants	2	
	Subalpine and alpine acidic	12,949	*
	Subalpine and alpine calcicolous grassy formations	3452	*
	Mountain	489	
	Arid and thermophilic	188	
	Megaforbieti of the alpine plain	11	

Habitats marked with \* are the ones used for this study.

A total of 336 soil samples were obtained from 122 plots. In each plot, the soil samples derived from three layers (0–10 cm; 10–20 cm; 20–40 cm) were extracted from three mini-pits (Figure S1); then, for each layer, a composite sample ( $\approx 100$  g) was created, mixing the samples derived from the three mini-pits. The sample totals and their related depths were as follows—122 samples taken from the 0–10 cm layer (layer I), 120 samples from the 10–20 cm layer (layer II), and 94 samples from 20–40 cm layer (layer III). Sampling by layers allowed comparisons among different plots that might have different soil layers.

To create the eDNA data, we used 156 samples (Table S2) of the 336 total samples. To perform the eDNA analysis, 20 g of soil were taken from layer I (80 samples) and layer II (76 samples). To avoid contamination, all eDNA sampling activities were performed with disposable gloves, and all tools were sterilized with a camping gas burner before each sampling. To limit the degradation of DNA during long-term storage, samples were inserted into sterile boxes (120 cc) with 40 g of Silica gel [26,27].

### 2.3. Environmental Features

Soil samples were analyzed to determine soil organic carbon (SOC), total nitrogen (TN), (Flash EA 1112 NCSoil, Thermo Fisher Scientific elemental analyzer, Pittsburgh, PA, USA), and pH in water (soil to water ratio of 1:2.5). Particle-size and distribution were also determined via sieving and sedimentation [28]. Finally, soil types were described according to the World Reference Base (WRB) [29].

Topsoils (or humus systems) were classified as Mull or Moder on the basis of the characteristics of OL (litter), OF (fragmented), OH (humus), and A (first mineral horizon) horizons [30].

Moreover, to make note of vegetation structure, all trees observed within the forest plots were recorded, and data on tree species and DBH (Diameter at Breast Height) were measured [31]. To study the forests, we determined indices of species and structural diversity following the advice of Chheng et al. [32]. Stem density (SD) was calculated as the number of trees per unit area (hectares). Basal area (BA) was measured as the cross-section of a tree's trunk measured at breast height (1.3 m) and then calculated for the whole plot. Lastly, the Shannon Index (HSH) was used as a diversity index of tree species [33].

#### 2.4. Molecular Analyses

Environmental DNA extraction was performed as follows: First, the 20 g of soil samples were mixed with 20 mL of phosphate buffer for 15 min [34]. Next, 2 mL of this mixture were centrifuged (10 min at 10,000 g) using the NucleoSpin<sup>®</sup> Soil Mini Kit (Macherey-Nagel, Düren, Germany). An amount of 400 µL of the resulting supernatant was kept as a starting material for the extraction, with a final elution of 150 µL. We also performed eight negative extraction controls, in which all extraction steps were performed without adding soil samples [22]. These controls were combined with the 156 soil samples for a total of 164 samples.

DNA amplification was performed for the following primer pairs: Arth02 [15] and Inse01 [15]. These pairs targeted arthropods and insects, respectively. Primers Arth02 and Inse01 amplify 16 S mitochondrial rDNA [15]. Forward and reverse primers were tagged on their 5'-end with an eight-nucleotide long sequence, in order to ensure a unique combination of primers for each sample. First, the 164 samples, 16 bioinformatic blanks, 2 positive controls, and 10 amplification controls were randomized on 96-well plates. Next, a qPCR assay was performed to determine the optimal number of amplification cycles, using 1 µL of 1:1000 diluted SYBR<sup>®</sup> Green I nucleic acid gel stain (Invitrogen<sup>™</sup>, Carlsbad, CA, USA), with undiluted and 1:10 diluted DNA. Bioinformatic blanks consisting of empty PCR-plate wells were used to monitor tag jumps, and the amplification controls were prepared by adding the amplification reagents to all, except to the ones used for DNA sample [16,35]. The positive control consisted of eDNA from a tropical forest soil provided by Pierre Taberlet. The positive control was used to monitor amplification and sequencing performance, and to also check for the cross-contamination of samples.

DNA amplification was performed using 2 µL of undiluted DNA, 2 µL of 5 µM forward and reverse primers, 0.16 µL of Bovine Serum Albumin (Roche Diagnostic, Basel, Switzerland), and 10 µL of AmpliTaq Gold 360 Master Mix 2 (Applied Biosystems<sup>™</sup>, Foster City, CA, USA). All reactions had a 20-µL volume, and four replicates per samples were performed in 384-well plates. Amplifications had an initial step of 10 min at 95 °C, and a denaturation step of 30 s at 95 °C. For 44 (Arth02) and 55 (Inse01), an annealing step of 30 s was performed either at 49 °C (Arth02) or 52 °C (Inse01), followed by an elongation step of 60 s at 72 °C. Subsequently, amplicons were visualized by high-resolution capillary electrophoresis (QIAxcel Advanced System; QIAGEN, Hilden, Germany), using 5 µL of pooled amplification product per marker. The purpose of these visualizations was to monitor both the expected fragment length (max. 170 bp for Arth02 and 270 bp for Inse01) and the primer dimers formation.

Amplicon purification was performed, per marker, as follows—(1) the four amplicon replicates per sample were pooled together, (2) six subsamples were purified using the MinElute PCR Purification Kit (QIAGEN, Hilden, Germany), and (3) the six subsamples were pooled together. Libraries were prepared following the MetaFast protocol [15] created by FASTERIS (Geneva, Switzerland), and the DNA was sequenced using the HiSeq 2500 (Arth02) and the NextSeq High (Inse01) (Illumina, San Diego, CA, USA).

#### 2.5. Bioinformatic Treatment of Sequence Data

Sequence data were analyzed using the OBITools software suite [35]. The analysis included the following steps: assembling of forward and reverse reads; using the *illumina-paired-end* program with an alignment score > 40; assigning sequences to the corresponding sample; using the *ngsfilter* program with zero and two mismatches allowed for the tags and primers, respectively; testing for sequence dereplication using the *obiuniq* program; filtering sequences with lengths shorter than 76 bp (Arth02) and 75 bp (Inse01), as well as singletons and low-quality sequences (i.e., sequences containing bases other than A, C, G, T); detection of PCR and sequencing errors using the *obiclean* program with the *-r* 0.5 option; constructing a reference database from EMBL (version 136); using the *ecoPCR* program [36]; and giving sequence taxonomical assignments using the *ecotag* program. Even though annelids were not a target of our primers, metabarcoding successfully amplified this taxon. Since many taxonomic



assignments of annelids were obtained, we included them in the study. However, we acknowledge that their representation is not complete.

Additional filtering of the taxonomically assigned sequences (MOTU) was performed in R (version 3.6.1; R Core Team, 2018), with the aim of removing MOTUs that can lead to biased results. Therefore, we removed MOTUs with a best identity value of  $<0.8$ ; MOTUs observed less than 10 times (Arth02) and 30 times (Inse01) in one sample; MOTUs observed more than one time in extraction or the amplification controls [22]; and MOTUs observed in only one out of the four PCR replicates [23].

### 3. Statistical Analysis

We used a canonical correlation analysis (CCA) to assess the relationships between pedofauna, habitat features, and the services provided by the ecosystems (e.g., carbon storage). CCA is a multivariate gradient analysis widely used to examine the relationship between environmental variables and community structure. This method measures the linear relationship between two multidimensional variables, and it is applied to the study of two groups of variables and their relations [37]. The output of a CCA is usually represented in figures with two-dimensional maps with vectors. The angles and axes of the maps represent the power of correlation (e.g., smaller angles represent a larger power), and the line length represents the strength of an environmental impact. Finally, the direction of the vectors relative to the axes determines whether the correlations are positive or negative. CCA assumes a unimodal response model of the variables that makes it possible to estimate the partial effect of variables and verify different relationships between the biotic and abiotic components of the environment. The statistical significance of the relationships between the explanatory and response variables was obtained with Montecarlo permutation procedures [38]. The results of the CCA are represented as biplots, where families of pedofauna and explanatory variables are projected on a bi-dimensional graph and defined by two canonical axes [39]. The statistical analyses were performed with the Canoco 4.5 software [39].

## 4. Results

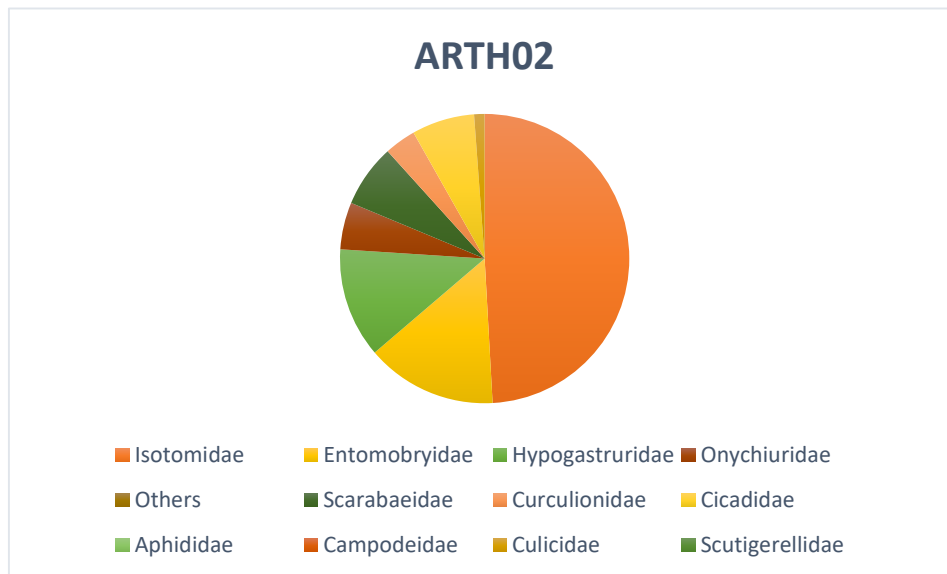
### 4.1. Soil Communities Described through eDNA Metabarcoding

Sequencing of the Arth02 and Inse01 libraries yielded a total of 12,545,971 and 25,966,237 raw sequences, which, after the bioinformatic treatment and filtering, consisted of 273 and 121 MOTUs, respectively.

The results of the eDNA analysis are presented according to the eDNA targets. Abundance was represented in terms of the relative frequency of sequences of a specific family found in an environmental sample.

Arthropods (Figure S2) detected by the Arth02 primer set were the first target. The most abundant arthropod family was Isotomidae (Figure 1). The composition of fauna community belonging to similar habitats, e.g., habitats that share environmental features (such as the mixed coniferous forest and larch forest), was fairly similar. These forests (Table 2) even contained the same collembolan families: Isotomidae, Entomobryidae, Hypogastruridae, and Onychiuridae. Spruce forest showed a similar pedofauna community, but compared to the larch and mixed coniferous forest, it had fewer sequences of Entomobryidae and Onychiuridae; then, the second most represented family in the spruce forest was Hypogastruridae. The two grasslands studied show similarities in community composition of arthropods, even though the second and the third most represented family differed from one other: in acidic grasslands were Cicadidae, Scarabaeidae, and Curculionidae, while in the calcicolous grasslands were Scarabaeidae, Hypogastruridae, and Curculionidae.

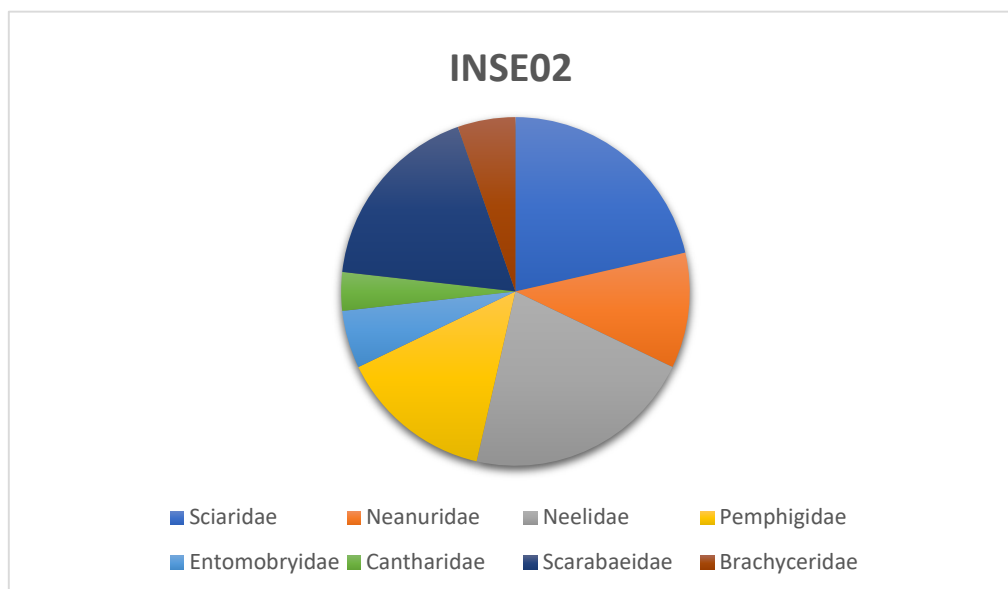




**Figure 1.** Representation of the Arthropods community read abundance detected with the primer ARTH02.

Interestingly, the mixed deciduous forest had a community composition different from the other habitats tested. It was mostly composed of Isotomidae, Cicadidae, and Culicidae.

Concerning the targets detected with the primer INSE02 (Table 2, Figure 2 and Figure S3), the most representative family detected depend on the habitat studied. For example, spruce forest and mixed coniferous forest had communities composed of mostly Pemphigidae, Neelidae, and Entomobryidae. Meanwhile, the larch forest showed a composition dominated by Sciaridae, Neanuridae, and four families with the same records: Neelidae, Pemphigidae, Entomobryidae, and Cantharidae.



**Figure 2.** Representation of the Arthropods community read abundance detected with the primer INSE02.

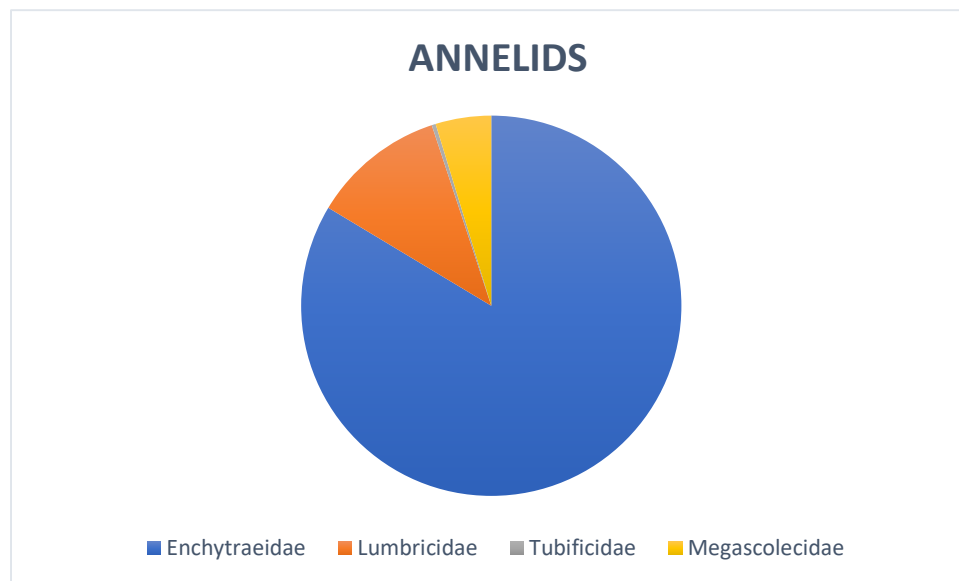
The grasslands also each had a different community composition. Indeed, even though the most represented family for both grasslands was Scarabaeidae, these habitats otherwise had quite different communities. The acidic prairie had a more heterogeneous composition and contained Sciaridae,



Neelidae, Brachyceridae, and Neanuridae, in addition to Scarabaeidae. The Calcicolous prairie only had Cantharidae, in addition to Scarabaeidae.

Mixed deciduous forest differed from every other habitat and had communities composed of Neelidae and Neanuridae.

When studying annelids, we detected that Enchytraeidae was the dominant group for each of the habitats examined (Table 3, Figure 3 and Figure S4). Lumbricidae was the second most common family in the spruce forest, larch forest, mixed coniferous forest, and calcicolous prairie. However, it was only the third most common family in the mixed deciduous forest and the acidic prairie, which both had higher levels of the Megascolecidae family.



**Figure 3.** Representation of Annelida community composition detected with eDNA metabarcoding.

**Table 3.** Presences of Annelida detected with eDNA metabarcoding.

Family Annelids	Larch Forest	Mixed Coniferous Forest	Mixed Deciduous Forest	Spruce Forest	Acidic Grassland	Calcicolous Grassland	Total
Enchytraeidae	104	93	18	35	217	28	495
Lumbricidae	27	14	3	5	17	1	67
Tubificidae	1	0	0	0	1	0	2
Megascolecidae	0	0	8	0	20	0	28

#### 4.2. Soil Communities and Environmental Features

The most commonly observed soil types were Cambisols, Regosols, Podzols, and Umbrisols. They were representative of the environments studied. For each soil type, values of pH in water ( $\text{pH}_{\text{H}_2\text{O}}$ ), soil skeleton levels, and soil organic carbon (SOC) levels were investigated. Soil analysis revealed that the Regosols had an average  $\text{pH}_{\text{H}_2\text{O}}$  value of 5.5 (min. 4.2, max 7.6). The rock fragment values for the Regosols ranged from an average of 35% in the first layer to 50% in the second layer and 60% in the third layer. The average SOC value for the Regosols was  $3.58 \text{ kg/m}^2$  (min. 0.82, max 9.39). For the Cambisols, the research team discovered an average  $\text{pH}_{\text{H}_2\text{O}}$  value of 5.5 (min. 4.4, max 7.4), rock fragments values ranged from an average of 20% in the first layer to 40% in the second layer and 50% in the third, and an average SOC level of  $6.27 \text{ kg/m}^2$  (min. 2.37, max 9.94). The Umbrisols had an average  $\text{pH}_{\text{H}_2\text{O}}$  value of 5.4 (min. 4.9 max 6.0), rock fragment values ranged from an average of 30% in the first layer to 55% in the second layer and 80% in the third, and an average SOC level of  $5.4 \text{ kg/m}^2$  (min. 1.87, max 12.39). The Podzols had an average  $\text{pH}_{\text{H}_2\text{O}}$  value of 4.9 (min. 4.4, max 6.1), rock fragment values ranged from an average of 25% in the first layer to 45% in the second layer and 60% in the third layer, and an average SOC level of  $4.12 \text{ kg/m}^2$  (min. 2.73, max 5.78).

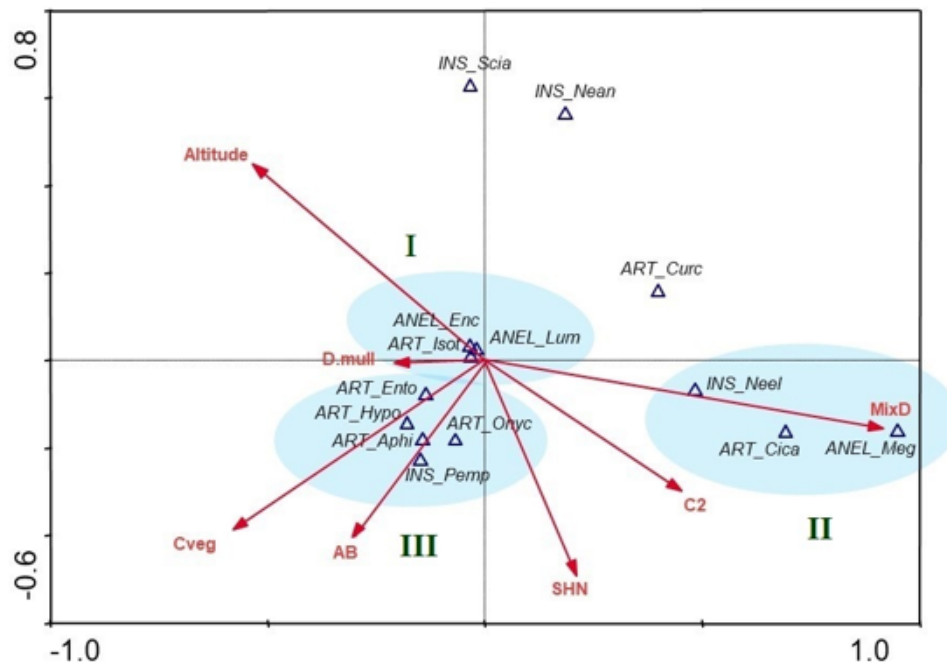
The statistical analyses were carried out separately for forests and grasslands as these environments were different enough that different environmental variables had to be considered (Table 4).

**Table 4.** Environmental features used in canonical correlation analysis (CCA) analysis.

Variable	Cod.	Units	Min	Max	Mean
Altitude	Altitude	m	744.00	2835.00	1904.24
Exposure	Esp	No unit			
Slope	Slope	Degree	0.00	48.00	20.68
Organic Carbon stock (0–10 cm)	C1	$\text{kg/m}^2$	0.48	8.21	3.19
Organic Carbon stock (10–20 cm)	C2	$\text{kg/m}^2$	0.09	4.57	1.48
pH (0–20 cm)	pH	No unit	4.05	7.70	5.42
Organic Carbon stock layer OH	OH	$\text{kg/m}^2$	0	1.707819	0
Organic Carbon stock layer OL and OF	OLOF	$\text{kg/m}^2$	0	2.658436	0.6
Organic Carbon stock of trees	Cveg	kg	105	15,986.93	77,019.73
Basal Area	AB	$\text{m}^2/\text{ha}$	0.53	68.78	37.38
Steam density	SD	Number of trees/ha	66.67	2611.11	760.28
Shannon index of trees	SHN	No unit	0.00	1.79	0.79
Acidic prairie	P_a	Presence/Absence			
Larch forest	Larix	Presence/Absence			
Mixed coniferous forest	MixC	Presence/Absence			
Mixed deciduous forest	MixD	Presence/Absence			
Cambisols	Cambi	Presence/Absence			
Podzols	Pdzl	Presence/Absence			
Regosols	Reg	Presence/Absence			
Umbrisols	Umbrisols	Presence/Absence			
Dysmoder	D.moder	Presence/Absence			
Dysmulls	D.mull	Presence/Absence			
Hemimoders	H.moder	Presence/Absence			
Mesomulls	M.mull	Presence/Absence			

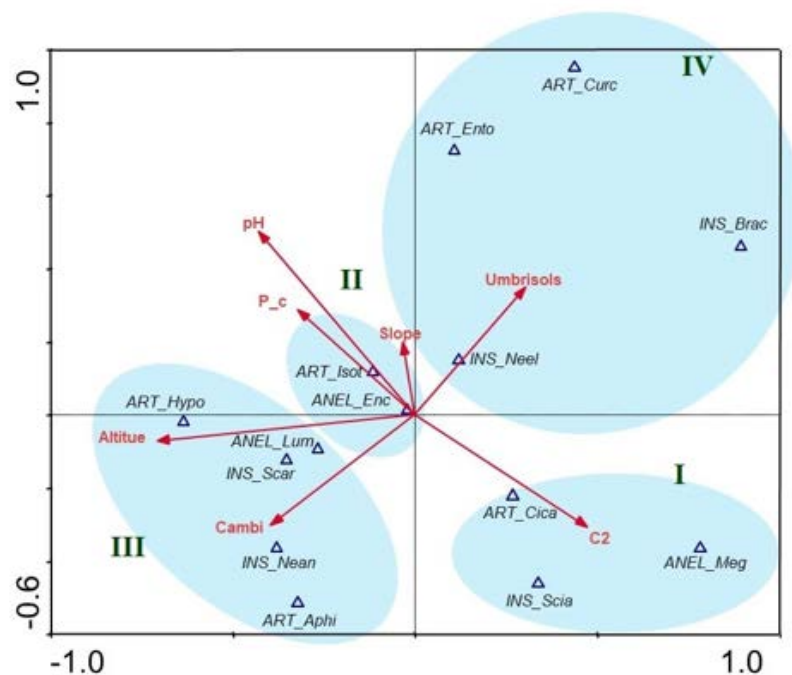
When studying the forests (Figure 4, Table S3), the cumulative variance of the species-environment relation was 31.4% for the first component, and 47.6% for the second. This explained almost 80% of the species-environment relation. The CCA studies of forests revealed three clusters. The first one, representing larch forest, correlated to higher altitudes and lower values of the Shannon index for vegetation. This was attributable to the fact that the larch forests were mostly composed of *Larix decidua* and no other tree species. The families that were related to this habitat were Sciaridae and Neanuridae, both belonging to the primer Inse02, in addition, the CCA analysis showed a weak correlation with other generalist families, such as Enchytraeidae and Isotomidae (Arth02). The second cluster was associated with the mixed deciduous forest, and the correlated families in this type of forest

were Megascolecidae for Annelida, Cicadidae for Arthropods (Arth02), and Neelidae for Arthropods (Inse02). The correlated environmental features for the mixed deciduous forest were lower altitude values and higher values of organic carbon stocked in the soil. Families belonging to this cluster also correlated with southern exposures, which suggest that there might be a relation with higher periods of exposure to direct sunlight [40]. This was also correlated with air and soil warming [41]. The third cluster occurred where mixed coniferous and spruce forests merged. The families here were the most generalist and had less specific environmental needs. The significantly related environmental features of these areas were the levels of organic carbon stocked in vegetation and AB values.



**Figure 4.** Biplot of forests environmental features and soil communities resulting from CCA analysis (environmental features are indicated in red, soil communities are indicated in black with a blue triangle, and blue circles with Roman numeral indicate clusters). The suffixes ART, ANEL, and INS refer to the targets: arthropods, annelids, and insects.

The CCA analysis of the grasslands explained more than 70% of the relation between species and environment (Figure 5, Table S4). In particular, the cumulative percentage variance of the species-environment relation was 26.8% for the first component and 45.7% for the second.



**Figure 5.** Biplot of grassland environmental features and soil communities resulting from CCA analysis (environmental features are indicated in red, soil communities are indicated in black with a blue triangle, and blue circles with the Roman numeral indicate clusters). The suffixes ART, ANEL, and INS refer to the targets: arthropods, annelids, and insects.

For the grasslands, the main identified ecological gradients were that of altitude, soil type, and those of the other environmental variables. Four clusters were identified.

The first cluster, representing acidic prairie, related to lower altitudes, southern exposures, and higher values of organic carbon stocked in soils. This cluster also correlated with more acidic pH values. The correlated families were Cicadidae for Arthropods (Arth02), Sciaridae for Arthropods (Inse02), and Megascolecidae for Annelida. The second cluster related to the environmental features present in the calcicolous prairie. This cluster related to the higher values of pH (according to this definition), and the families belonging to this cluster were the most generalist, had less specific environmental needs, and included Isotomidae for arthropods and Enchytraeidae for Annelida. The remaining clusters were related more to the environmental features, such as soil types. The third cluster was related to Cambisols and higher altitudes, with a heterogeneous composition of pedofauna, including Lumbricidae, Hypogastruridae, Neanuridae, and Aphidae. While, the fourth cluster was related to the Umbrisols and families, including Neelidae, Entomobryidae, and Brachymeridae.

## 5. Discussion

The data obtained using eDNA metabarcoding showed similarities in the community composition of arthropods and annelids among similar habitats, such as spruce forest and mixed coniferous forest. Nevertheless, the analysis revealed that habitats that were quite different, such as grasslands and mixed coniferous forests, had notably different community compositions. With regard to the mixed deciduous forest, this forest habitat showed a peculiar pedofauna community composition. In some instances, this could be related to the characteristics of the acidic prairie and its annelids. Further investigation is needed to determine if there are any environmental features that these habitats share and their relationship with the current pedofauna.

Arthropods are related to higher values of heterogeneity in family composition in habitats with coniferous vegetation, and annelids are more heterogeneous in the acidic prairie and the mixed deciduous forest.

Some families are quite common among all habitats and are easily found during investigations, thus, they could be described as generalists without specific environmental needs. They might also be considered to be families composed of a numerous genus, such as Enchytraeidae [42] for annelids or Isotomidae for arthropods. For this reason, they can be found almost everywhere. Other families, such as Cicadidae for Arthropods (Inse02) or Megascolecidae for annelids, are more selective of their habitats and one can theorize that they probably have more specific needs than the generalist families. Populations of generalist species are more likely to adapt to new habitats [43,44], and generalist species are less likely to be endangered. On the other hand, species with specialized niche and environmental needs might struggle to thrive in a new habitat [45] or to cope with changes driven by climate change or habitat management. Further research must be done in order to understand if these families should be subject to environmental protection. Indeed, even families recorded in many habitats might still contain endangered species. In general, the arthropod taxon requires much additional research.

The canonical correlation analysis showed that there were differences between grasslands and forests regarding pedofauna composition. This was an expected result that confirmed the results of the eDNA analysis. However, the distinctions between the compositions of these environments were not very clear, due to the fact that the taxonomic resolution—at least of this study—Only resolved to the family level. Environmental DNA metabarcoding is a relatively new research technique. Sometimes a better taxonomic resolution, such as species resolution, is not achievable due to incomplete reference databases, and many sequences are yet to be filed. This is true above all for the sequences of endemic species varieties, such as alpine fauna.

Not being able to do research at the species level might have led to the loss of important information regarding ecological needs and their relations to environmental features. Therefore, the ecological descriptions yielded by this study are generalized.

Despite the limitation of taxonomic resolution through eDNA metabarcoding, we described the edaphic soil communities over wide study areas, covering the most representative habitat in an alpine area. Additionally, traditional studies about pedofauna were carried out in the same study area by the National Park [46], detecting four families and reaching the taxonomic resolution of species, but covering only few plots. Thus, these two techniques were not compared, rather they could be integrated in order to yield an in-depth description of soil biodiversity in specific areas, integrating data from the pitfall traps that mostly describe surface pedofauna and data from eDNA metabarcoding that describe pedofauna at multiple levels of depth in the soil.

It is important to study the biodiversity of pedofauna because biodiversity data in terms of taxonomic descriptions, ecological roles, and sensitivity to changes are still poor [6]. The taxa studied here play important roles in soil functioning. They influence the detritus food chain, organic matter degradation, pollination, pest control, soil aeration, and soil turnover [47–49].

For this reason, it would be interesting to improve studies of these taxa and to develop different approaches that could complement or improve upon the traditional taxonomy. For example, studies on the ecological role of families or communities would make it possible to understand in a functional way the role that these families play in soil and the functions of the ecosystem to which they belong. Some studies have already classified invertebrates using functional classifications, such as trophic classifications, based on their role in the food webs of organisms [50,51] or classifications based on the relation between invertebrate attributes and environmental features [52]. If a better understanding of ecosystem functions can be developed, there will be a corresponding simplification when it comes to evaluating the associated ecosystem services. This study attempted to take a first step in this direction by studying the community composition of soil and the related features between communities and their environments. Still, our understanding of the correlation of cause-effect between families and environmental features is lacking, and further research must be done.

## 6. Conclusions

The results of this study demonstrated the complexity of soil fauna communities by characterizing them habitat-to-habitat in the study area. This study, by making a comparison between habitats, also revealed that changes across habitats are due to various environmental needs. This study was a first step in a broader effort aimed at increasing our comprehension of the complex interaction of the ecological roles and ecosystem functions being performed every day in the soils of the Gran Paradiso National Park. Once these interactions are understood, ecosystem services related to biodiversity can be properly evaluated, and adequate and appropriate conservation actions can be taken.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1999-4907/11/7/738/s1>, Figure S1: Soil sampling scheme, Figure S2: Community composition of Arthropods (primer ARTH02) for each habitat investigated with eDNA metabarcoding, Figure S3: Community composition of Arthropods (primer INSE02) for each habitat investigated with eDNA metabarcoding, Figure S4: Community composition of Anellida for each habitat investigated with eDNA metabarcoding, Table S1: Distribution of plots per habitat type, Table S2: Distribution of eDNA samples per layer, Table S3: Effects of environmental features of forests in the CCA analysis (\*\* =  $p < 0.05$ ), Table S4: Effects of environmental features of grasslands in the CCA analysis (\*\* =  $p < 0.05$ ; \* =  $p < 0.1$ ).

**Author Contributions:** Conceptualization, C.C., E.P.-S., F.F., and N.R.; methodology, C.C., E.P.-S., C.F., and G.F.F.; software, N.R., C.C., C.F., A.G., and G.F.F.; formal analysis, N.R., C.C., C.F., A.G., and G.F.F.; data curation, N.R., C.C., C.F., and A.G.; writing—Original draft preparation, N.R.; writing—Review and editing, C.C., E.P.-S., C.F., A.G., and G.F.F.; supervision, C.C. and E.P.-S.; project administration, E.P.-S. and C.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was founded with a Ministerial grant (Ministry of Education, University and Research cod. 18A1/11) and by the European Research Council under the European Community's Horizon 2020 Programme, Grant Agreement no. 772284 (IceCommunities).

**Acknowledgments:** We thank the Gran Paradiso National Park for providing essential assistance to conduct the field work for this research, in particular we want to thank A. Mingozi, B. Bassano, and R. Viterbi. A.G. thanks Aurélie Bonin for the constructive discussions during the bioinformatic treatments of data. We also thank the two anonymous reviewers that helped to improve our manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Lemanceau, P. EcoFINDERS characterize biodiversity and the function of soils in Europe 23 partners in 10 European countries and China. *Biofutur* **2011**, *326*, 56–58.
2. Prather, C.M.; Pelini, S.L.; Laws, A.; Rivest, E.; Woltz, M.; Bloch, C.P.; Del Toro, I.; Ho, C.K.; Kominoski, J.; Newbold, T.A.S.; et al. Invertebrates, ecosystem services and climate change. *Biol. Rev.* **2013**, *88*, 327–348. [[CrossRef](#)] [[PubMed](#)]
3. Plewa, R.; Jaworski, T.; Hilszczański, J.; Horák, J. Investigating the biodiversity of the forest strata: The importance of vertical stratification to the activity and development of saproxylic beetles in managed temperate deciduous forests. *For. Ecol. Manag.* **2017**, *402*, 186–193. [[CrossRef](#)]
4. Kopali, A.; Gixhari, B.; Vrapı, H.; Ismaili, H.; Uka, R. Impact of biodiversity (pedo-fauna) in the biological qualities of soil in biological and conventional agro-ecosystems, in apple tree (*Malus domestica*). *Online Int. Interdiscip. Res. J.* **2014**, *4*, 61–69.
5. Seastedt, T.R. The role of microarthropods in decomposition and mineralization processes. *Annu. Rev. Entomol.* **1984**, *29*, 25–46. [[CrossRef](#)]
6. Nielsen, U.N.; Ayres, E.; Wall, D.H.; Bardgett, R.D. Soil biodiversity and carbon cycling: A review and synthesis of studies examining diversity-function relationships. *Eur. J. Soil Sci.* **2011**, *62*, 105–116. [[CrossRef](#)]
7. Falco, L.; Coviella, C. Contribution of *Amyntas gracilis* (Megascolecidae) and *Octolasion cyaneum* (Lumbricidae) to soil physical stability: A mesocosm experiment. *Rev. Fac. Cienc. Agrar.* **2016**, *48*, 115–127.
8. van Vliet, P.C.J.; West, L.T.; Hendrix, P.F.; Coleman, D. The influence of Enchytraeidae (Oligochaeta) on the soil porosity of small microcosms. *Geoderma* **1993**, *56*, 287–299. [[CrossRef](#)]
9. Rumpel, C.; Kögel-Knabner, I. Deep soil organic matter—a key but poorly understood component of terrestrial C cycle. *Plant Soil* **2011**, *338*, 143–158. [[CrossRef](#)]



10. Lefèvre, C.; Rekik, F.; Alcantara, V.; Wiese, L. *Soil Organic Carbon: The Hidden Potential*; Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 2017; ISBN 9251096813.
11. Wiesmeier, M.; Urbanski, L.; Hobley, E.; Lang, B.; von Lützow, M.; Marin-Spiotta, E.; Wollschläger, U. Soil organic carbon storage as a key function of soils—A review of drivers and indicators at various scales. *Geoderma* **2019**, *333*, 149–162. [[CrossRef](#)]
12. Hasegawa, M. The relationship between the organic matter composition of a forest floor and the structure of a soil arthropod community. *Eur. J. Soil Biol.* **2001**, *37*, 281–284. [[CrossRef](#)]
13. Lavelle, P. Diversity of soil fauna and ecosystem function. *Biol. Int.* **1996**, *33*, 3–16.
14. Naeem, S.; Thompson, L.J.; Lawler, S.P.; Lawton, J.H.; Woodfin, R.M. Empirical evidence that declining species diversity may alter the performance of terrestrial ecosystems. *Philos. Trans. R. Soc. B Biol. Sci.* **1995**, *347*, 249–262. [[CrossRef](#)]
15. Taberlet, P.; Bonin, A.; Coissac, E.; Zinger, L. *Environmental DNA: For Biodiversity Research and Monitoring*; Oxford University Press: Oxford, UK, 2018; ISBN 0198767226.
16. Welch, D.B.M.; Huse, S.M. Microbial Diversity in the Deep Sea and the Underexplored “Rare Biosphere”. *Handb. Mol. Microb. Ecol. II Metagenom. Differ. Habitats* **2011**, *103*, 243–252. [[CrossRef](#)]
17. Gómez-Brandón, M.; Ascher-Jenuell, J.; Bardelli, T.; Fornasier, F.; Sartori, G.; Pietramellara, G.; Arfaioli, P.; Egli, M.; Beylich, A.; Insam, H.; et al. Ground cover and slope exposure effects on micro- and mesobiota in forest soils. *Ecol. Indic.* **2017**, *80*, 174–185. [[CrossRef](#)]
18. Yang, C.; Wang, X.; Miller, J.A.; De Blécourt, M.; Ji, Y.; Yang, C.; Harrison, R.D.; Yu, D.W. Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. *Ecol. Indic.* **2014**, *46*, 379–389. [[CrossRef](#)]
19. Baird, D.J.; Hajibabaei, M. Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Mol. Ecol.* **2012**, *21*, 2039–2044. [[CrossRef](#)]
20. Lear, G.; Dickie, I.; Banks, J.; Boyer, S.; Buckley, H.; Buckley, T.; Cruickshank, R.; Dopheide, A.; Handley, K.; Hermans, S.; et al. Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. *N. Z. J. Ecol.* **2018**, *42*, 10–51. [[CrossRef](#)]
21. Ficetola, G.F.; Taberlet, P.; Coissac, E. How to limit false positives in environmental DNA and metabarcoding? *Mol. Ecol. Resour.* **2016**, *16*, 604–607. [[CrossRef](#)] [[PubMed](#)]
22. Zinger, L.; Bonin, A.; Alsos, I.G.; Bálint, M.; Bik, H.; Boyer, F.; Chariton, A.A.; Creer, S.; Coissac, E.; Deagle, B.E. DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Mol. Ecol.* **2019**, *28*, 1857–1862. [[CrossRef](#)]
23. Ficetola, G.F.; Pansu, J.; Bonin, A.; Coissac, E.; Giguët-Covex, C.; De Barba, M.; Gielly, L.; Lopes, C.M.; Boyer, F.; Pompanon, F.; et al. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Mol. Ecol. Resour.* **2015**, *15*, 543–556. [[CrossRef](#)] [[PubMed](#)]
24. Klanderud, K.; Totland, Ø. Simulated climate change altered dominance hierarchies and diversity of an alpine biodiversity hotspot. *Ecology* **2005**, *86*, 2047–2054. [[CrossRef](#)]
25. Courbaud, B.; Kunstler, G.; Morin, X.; Cordonnier, T. What is the future of the ecosystem services of the Alpine forest against a backdrop of climate change? *Rev. Géographie Alp.* **2011**, 1–12. [[CrossRef](#)]
26. Chase, M.W.; Hills, H.H. Silica gel: An ideal material for field preservation of leaf samples for DNA studies. *Taxon* **1991**, *40*, 215–220. [[CrossRef](#)]
27. Guo, Y.; Yang, G.; Chen, Y.; Li, D.; Guo, Z. A comparison of different methods for preserving plant molecular materials and the effect of degraded DNA on ddRAD sequencing. *Plant Divers.* **2018**, *40*, 106–116. [[CrossRef](#)]
28. Burt, R. *Soil Survey Laboratory Methods Manual*; Soil Survey Investigations Report No. 42, Version 4.0; USDA-NRCA: Washington, DC, USA, 2004.
29. IUSS Working Group WRB. *World Reference Base for Soil Resources 2014, Update 2015*; World Soil Resources Reports 106; FAO: Rome, Italy, 2015.
30. Zanella, A.; Ponge, J.F.; Jabiol, B.; Sartori, G.; Kolb, E.; Le Bayon, R.C.; Gobat, J.M.; Aubert, M.; De Waal, R.; Van Delft, B.; et al. Humusica 1, article 5: Terrestrial humus systems and forms—Keys of classification of humus systems and forms. *Appl. Soil Ecol.* **2018**, *122*, 75–86. [[CrossRef](#)]
31. Wang, C. Biomass allometric equations for 10 co-occurring tree species in Chinese temperate forests. *For. Ecol. Manag.* **2006**, *222*, 9–16. [[CrossRef](#)]
32. Chheng, K.; Sasaki, N.; Mizoue, N.; Khorn, S.; Kao, D.; Lowe, A. Assessment of carbon stocks of semi-evergreen forests in Cambodia. *Glob. Ecol. Conserv.* **2016**, *5*, 34–47. [[CrossRef](#)]

33. Shannon, C.E.; Weaver, W. *A Mathematical Model of Communication*; University of Illinois Press: Urbana, IL, USA, 1949; Volume 11, pp. 1–117.
34. Taberlet, P.; Coissac, E.; Pompanon, F.; Brochmann, C.; Willerslev, E. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol. Ecol.* **2012**, *21*, 2045–2050. [[CrossRef](#)]
35. Boyer, F.; Mercier, C.; Bonin, A.; Le Bras, Y.; Taberlet, P.; Coissac, E. obitools: A unix-inspired software package for DNA metabarcoding. *Mol. Ecol. Resour.* **2016**, *16*, 176–182. [[CrossRef](#)] [[PubMed](#)]
36. Ficetola, G.F.; Coissac, E.; Zundel, S.; Riaz, T.; Shehzad, W.; Bessière, J.; Taberlet, P.; Pompanon, F. An in silico approach for the evaluation of DNA barcodes. *BMC Genom.* **2010**, *11*, 434. [[CrossRef](#)] [[PubMed](#)]
37. Ter Braak, C.J.F. Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology* **1986**, *67*, 1167–1179. [[CrossRef](#)]
38. Lepš, J.; Šmilauer, P. *Multivariate Analysis of Ecological Data Using CANOCO*; Cambridge University Press: Cambridge, UK, 2003; ISBN 0521891086.
39. Ter Braak, C.J.F.; Šmilauer, P. *CANOCO Reference Manual and CanoDraw for Windows User's Guide. Software Canonical Community Ordination (Version 4.5)*; Microcomputer Power: Ithaca, NY, USA, 2002.
40. Salmon, S. Changes in humus forms, soil invertebrate communities and soil functioning with forest dynamics. *Appl. Soil Ecol.* **2018**, *123*, 345–354. [[CrossRef](#)]
41. Imbeck, H.; Ott, E. Ecology and regeneration of a subalpine Norway spruce forest rich in tall forbs, considering especially snow and avalanche problems. *Mitt. Eidg. Inst. Schnee Lawinenforsch* **1987**, *42*, 1–202.
42. Raefe, U.G.; Chmelz, R.M.S. Indicator values, strategy types and life forms of terrestrial Enchytraeidae and other microannelids. *Newsl. Enchytraeidae* **1999**, *6*, 59–67.
43. Angert, A.L.; Crozier, L.G.; Rissler, L.J.; Gilman, S.E.; Tewksbury, J.J.; Chuncó, A.J. Do species' traits predict recent shifts at expanding range edges? *Ecol. Lett.* **2011**, *14*, 677–689. [[CrossRef](#)] [[PubMed](#)]
44. Warren, M.S.; Hill, J.K.; Thomas, J.A.; Asher, J.; Fox, R.; Huntley, B.; Roy, D.B.; Telfer, M.G.; Jeffcoate, S.; Harding, P.; et al. Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature* **2001**, *414*, 65–69. [[CrossRef](#)]
45. Gilman, S.E.; Wethey, D.S.; Helmuth, B. Variation in the sensitivity of organismal body temperature to climate change over local and geographic scales. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9560–9565. [[CrossRef](#)] [[PubMed](#)]
46. Viterbi, R.; Cerrato, C.; Bassano, B.; Bionda, R.; Hardenberg, A.; Provenzale, A.; Bogliani, G. Patterns of biodiversity in the northwestern Italian Alps: A multi-taxa approach. *Community Ecol.* **2013**, *14*, 18–30. [[CrossRef](#)]
47. Grêt-Regamey, A.; Brunner, S.H.; Kienast, F. Mountain Ecosystem Services: Who Cares? *Mt. Res. Dev.* **2012**, *32*, S23–S34. [[CrossRef](#)]
48. Barbati, A.; Ferrari, B.; Alivernini, A.; Quatrini, A.; Merlini, P.; Puletti, N.; Corona, P. Forest ecosystems and carbon sequestration in Italy. *L'italia For. e Mont.* **2014**, *69*, 205–212. [[CrossRef](#)]
49. de Vasconcellos, R.C.; Beltrão, N.E.S. Avaliação de prestação de serviços ecossistêmicos em sistemas agroflorestais através de indicadores ambientais. *Interações* **2018**, *19*, 209. [[CrossRef](#)]
50. Saigo, M.; Marchese, M.; Wantzen, K.M. A closer look at the main actors of neotropical floodplain food webs: Functional classification and niche overlap of dominant benthic invertebrates in a floodplain lake of paraná river. *Iheringia Ser. Zool.* **2016**, *106*, 1–8. [[CrossRef](#)]
51. Cummins, K.W.; Merritt, R.W.; Andrade, P.C.N. The use of invertebrate functional groups to characterize ecosystem attributes in selected streams and rivers in south Brazil. *Stud. Neotrop. Fauna Environ.* **2005**, *40*, 69–89. [[CrossRef](#)]
52. Oliveira, P.H.G.; Gama-Rodrigues, A.C.; Gama-Rodrigues, E.F.; Sales, M.V.S. Litter and soil-related variation in functional group abundances in cacao agroforests using structural equation modeling. *Ecol. Indic.* **2018**, *84*, 254–262. [[CrossRef](#)]

