

# Metabarcoding data reveal vertical multi-taxa variation in topsoil communities during the colonization of deglaciated forelands

Alessia Guerrieri<sup>1</sup>, Alexis Carteron<sup>1</sup>, Aurélie Bonin<sup>2</sup>, Silvio Marta<sup>1</sup>, Roberto Ambrosini<sup>1</sup>, Marco Caccianiga<sup>1</sup>, Isabel Cantera<sup>1</sup>, Chiara Compostella<sup>1</sup>, Guglielmina Diolaiuti<sup>1</sup>, Diego Fontaneto<sup>3</sup>, Ludovic Gielly<sup>4</sup>, Fabrizio Gili<sup>1</sup>, Mauro Gobbi<sup>5</sup>, Jérôme Poulenard<sup>6</sup>, Pierre Taberlet<sup>4</sup>, Andrea Zerboni<sup>1</sup>, Wilfried Thuiller<sup>4</sup>, and Gentile Francesco Ficetola<sup>1</sup>

<sup>1</sup>Università degli Studi di Milano

<sup>2</sup>Argaly

<sup>3</sup>Water Research Institute National Research Council Verbania Branch

<sup>4</sup>Universite Grenoble Alpes

<sup>5</sup>MUSE-Science Museum trento

<sup>6</sup>Université Savoie Mont-Blanc

January 18, 2022

## Abstract

Ice-free areas are increasing worldwide due to the dramatic glacier shrinkage and are undergoing rapid colonization by multiple lifeforms, thus representing key environments to study ecosystem development. Soils have a complex vertical structure. However, we know little about how microbial and animal communities differ across soil depths and development stages during the colonization of deglaciated terrains, how these differences evolve through time, and whether patterns are consistent among different taxonomic groups. Here, we used environmental DNA metabarcoding to describe how community diversity and composition of six groups (Eukaryota, Bacteria, Mycota, Collembola, Insecta, Oligochaeta) differ between surface (0-5 cm) and relatively deep (7.5-20 cm) soils at different stages of development across five Alpine glaciers. Taxonomic diversity increased with time since glacier retreat and with soil evolution; the pattern was consistent across different groups and soil depths. For Eukaryota, and particularly Mycota, alpha-diversity was generally the highest in soils close to the surface. Time since glacier retreat was a more important driver of community composition compared to soil depth; for nearly all the taxa, differences in community composition between surface and deep soils decreased with time since glacier retreat, suggesting that the development of soil and/or of vegetation tends to homogenize the first 20 cm of soil through time. Within both Bacteria and Mycota, several molecular operational taxonomic units were significant indicators of specific depths and/or soil development stages, confirming the strong functional variation of microbial communities through time and depth. The complexity of community patterns highlights the importance of integrating information from multiple taxonomic groups to unravel community variation in response to ongoing global changes.

## Metabarcoding data reveal vertical multi-taxa variation in topsoil communities during the colonization of deglaciated forelands

Running title: Multi-taxa variation of topsoil communities

Alessia Guerrieri<sup>1</sup>, Alexis Carteron<sup>1,\*</sup>, Aurélie Bonin<sup>1,2</sup>, Silvio Marta, Roberto Ambrosini<sup>1</sup>, Marco Caccianiga<sup>1</sup>, Isabel Cantera<sup>1</sup>, Chiara Compostella<sup>3</sup>, Guglielmina Diolaiuti<sup>1</sup>, Diego Fontaneto<sup>4</sup>, Ludovic Gielly<sup>5</sup>, Fabrizio Gili<sup>1</sup>, Mauro Gobbi<sup>6</sup>, Jerome Poulenard<sup>7</sup>, Pierre Taberlet<sup>5,8</sup>, Andrea Zerboni<sup>3</sup>, Wilfried Thuiller<sup>5</sup>, Gentile Francesco Ficetola<sup>1,5</sup>

Affiliations:

<sup>1</sup>Department of Environmental Science and Policy, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy

<sup>2</sup>Argaly, Bâtiment CleanSpace, 354 Voie Magellan, 73800 Sainte-Hélène-du-Lac, France

<sup>3</sup>Department of Earth Sciences “A. Desio”, Università degli Studi di Milano, Via L. Mangiagalli 34, 20133 Milano, Italy

<sup>4</sup>CNR - Water Research Institute, Largo V. Tonolli 50, 28922, Verbania, Italy

<sup>5</sup>Univ. Grenoble Alpes, CNRS, Univ. Savoie Mont Blanc, LECA, Laboratoire d’Ecologie Alpine, F-38000, Grenoble, France

<sup>7</sup>Univ. Savoie Mont Blanc, CNRS, EDYTEM, F-73000 Chambéry

<sup>8</sup>UiT – The Arctic University of Norway, Tromsø Museum, Tromsø, Norway.

\*Correspondence author. Email: alexis.carteron@unimi.it

## ABSTRACT

Ice-free areas are increasing worldwide due to the dramatic glacier shrinkage and are undergoing rapid colonization by multiple lifeforms, thus representing key environments to study ecosystem development. Soils have a complex vertical structure. However, we know little about how microbial and animal communities differ across soil depths and development stages during the colonization of deglaciated terrains, how these differences evolve through time, and whether patterns are consistent among different taxonomic groups. Here, we used environmental DNA metabarcoding to describe how community diversity and composition of six groups (Eukaryota, Bacteria, Mycota, Collembola, Insecta, Oligochaeta) differ between surface (0-5 cm) and relatively deep (7.5-20 cm) soils at different stages of development across five Alpine glaciers. Taxonomic diversity increased with time since glacier retreat and with soil evolution; the pattern was consistent across different groups and soil depths. For Eukaryota, and particularly Mycota, alpha-diversity was generally the highest in soils close to the surface. Time since glacier retreat was a more important driver of community composition compared to soil depth; for nearly all the taxa, differences in community composition between surface and deep soils decreased with time since glacier retreat, suggesting that the development of soil and/or of vegetation tends to homogenize the first 20 cm of soil through time. Within both Bacteria and Mycota, several molecular operational taxonomic units were significant indicators of specific depths and/or soil development stages, confirming the strong functional variation of microbial communities through time and depth. The complexity of community patterns highlights the importance of integrating information from multiple taxonomic groups to unravel community variation in response to ongoing global changes.

**KEYWORDS** : environmental DNA, glacier retreat, Hill’s number, beta-diversity, soil depth, springtails, earthworms, insects, fungi

The worldwide shrinkage of glaciers is causing a fast increase in ice-free areas on all continents, thus habitats multiple organisms (Ficetola et al., 2021). After ice loss, organisms with high dispersal abilities can colonize the newly exposed terrains relatively quickly (Gobbi et al., 2017; Hågvar et al., 2020; Kaufmann, 2001; Rosero et al., 2021). Both micro- and macro-organisms (e.g. bacteria, fungi and soil fauna) influence soil development being involved in many biogeochemical processes such as soil nutrient cycling (Bardgett 2005; Bardgett & Van Der Putten, 2014), and can interact with each other in determining ecosystem functioning (Ingham, Trofymow, Ingham, & Coleman, 1985). Assessing community variation in this biodiversity compartment and across multiple glacier forelands is important to understand how these ecosystems develop after the retreat of glaciers, and is a key topic of global change biology (Cauvy-Fraunié & Dangles, 2019).

To date, most of our knowledge about soil ecology focuses on the top 10 cm of soil and on microbial communities (Bahram et al., 2015). However, soil characteristics vary vertically (Khokon, Schneider, Daniel, & Polle, 2021). In particular, physical features (e.g., pH, soil moisture, micro-climatic characteristics) typically

change through space, and the availability of nutrients (e.g., organic carbon, total nitrogen), together with the enzyme activities of the associated microorganisms, decrease from the topsoil to deeper soil layers (Herold et al., 2014; Moradi et al., 2020). Such variation of habitat conditions can strongly influence the community structure of inhabiting taxa (Carteron, Beigas, Joly, Turner, & Laliberté, 2021; Franzetti et al., 2020; Orwin, Kirschbaum, John, & Dickie, 2011; Rime et al., 2015) because different organisms can be associated with different soil conditions (Khokon, Schneider, Daniel, & Polle, 2021; Mundra et al., 2021). Several studies showed differences in microbial communities across soil depths in several terrestrial habitats including grasslands, forests, high elevation, post-mining and reforested-soils, and agreed that depth significantly affects the abundance, composition and diversity of bacteria and fungi, with the richest communities often associated to surface layers (e.g. Zhao, Zheng, Zhang, Gao, & Fan, 2021; Carteron et al., 2021; Chu et al., 2016; Moradi et al., 2020; Chen, Jiao, Li, & Du, 2020). These differences can be attributable to the decrease in nutrients content at increasing depths (Chu et al., 2016), or to differences in microclimatic conditions and water availability. However, the spatial structuring and micro-habitats conditions of soil communities are yet poorly known, and most studies only focused on very few limited taxonomic groups (Doblas-Miranda, Sánchez-Piñero, & González-Megías, 2009; Moradi et al., 2020; Sadaka & Ponge, 2003), making it difficult to compare the responses of functionally different taxa.

During soil formation after the retreat of glaciers, many features of the substrate change through time, with modifications of physical properties and nutrients content, and a progressive vertical stratification of developed soils (Schaetzl & Anderson, 2005; Mavris, Egli, Plotze, Blum, Mirabella and Giaccari, 2010; Khedim et al., 2021; Wietrzyk-Pełka, Rola, Szymański, & Węgrzyn, 2020). Despite many studies investigating the biotic colonization after glacier retreat, the majority of them focused on organisms living above or just below the surface (reviewed in Ficetola et al. 2021), while limited information is available about the vertical distribution of different topsoil organisms across stages of soil development. Assessing the vertical as well as the horizontal composition and distribution of topsoil colonizers along the glacier forelands is pivotal to infer the key ecological processes under the primary succession that occurs since the early years after glacier retreat. Rime et al. (2015) performed a rare attempt of integrating soil depth into the study of Alpine primary successions (see also Bajerski & Wagner, 2013; Schütte et al., 2009). They assessed the structure of microbial communities along one glacier foreland, and found that soil depth and development stage interact in shaping the biodiversity of bacteria and fungi. Differences between communities from surface and deep layers were particularly strong immediately after glacier retreat, while decreased at older soil development stages, with a homogenization through time. However, Rime et al. (2015) only focused on microorganisms and considered just one glacier foreland. As different topsoil organisms can have very different responses (Cauvy-Fraunié & Dangles, 2019; Donald et al., 2021; Ficetola et al., 2021; Rosero et al., 2021), the study of multiple taxa is needed for a better understanding of the ecological processing governing community development after glacier retreat.

Approaches based on the metabarcoding of environmental DNA (eDNA; Taberlet, Bonin, Zinger, & Coissac, 2018) help overcoming several limitations of conventional sampling and are increasingly used because of their relatively fast and cost-efficient data production. Environmental DNA metabarcoding allows the monitoring of communities of micro- and macro-organisms in a wide range of natural systems (Bohmann et al., 2014). With appropriate technical precautions (Guerrieri et al., 2020), soil communities can be sampled and studied via metabarcoding over broad geographic scales and from remote areas (e.g. Zinger, Taberlet, et al., 2019), and data can be related to environmental characteristics in order to infer ecological processes. The combination of multiple metabarcodes makes eDNA particularly powerful tool for estimating the multi-taxa soil diversity (Donald et al., 2021). Here, we used metabarcoding data from soil eDNA in order to study the vertical distribution of microbes and animals within the top 20 cm of soil, where most microbial diversity has been retrieved (Fierer, Schimel, & Holden, 2003) and where most soil invertebrates spend their life cycle (Menta, 2012). First, we tested whether and how the overall taxonomic diversity of multiple taxa changes with soil depth and time since glacier retreat. We expected that the alpha-diversity of communities increases through time and decreases with depth, especially in the youngest soils. Moreover, we tested whether the changes in alpha-diversity through time are consistent between surface and deep layers. Second,

we evaluated the differences in community composition between different depths and tested for potential taxa characteristic of the different depths or stages of soil development. Rime et al (2015) observed that differences between surface and deep soil decrease at older soil development stages, with a homogenization of communities through time, but these conclusions were only based on microorganisms from one single glacier foreland. We analyzed the beta-diversity between surface and deep layers for six taxonomic groups representing a large proportion of biodiversity. If the Rime’s homogenization hypothesis applies to the whole biota, we expect that beta-diversity between surface and deep layers decreases from recent to more developed terrains, with a consistent pattern across taxa.

## MATERIALS AND METHODS

### Samples collection and preservation

In Summer 2018, we collected 280 soil samples from five Alpine forelands (Fig. 1): Amola (coordinates of the center of the foreland: N 46.215° E 10.697°), Morteratsch (N 46.438° E 9.936°), Rutor (N 46.669° E 6.992°), Sforzellina (N 46.351° E 10.510°) and Grande di Verra (N 45.895° E 7.749°). For each foreland, we selected three to eight sites for which the date of glacier retreat is known on the basis of the literature, dated images and field surveys, focusing on the period between the end of the Little Ice Age (~1850) and recent years (Marta et al., 2021). Soil samples were representative of different stages of soil development depending on the time elapsed between the retreat of glaciers and sampling activities (hereafter referred to as “time since glacier retreat”; ranging from 12 to 168 years).

At each site, we established five regularly spaced plots at distances of about 20 m. At each plot, we collected five soil cores within one-meter distance and we kept the 0-5 cm and 7.5-20 cm portions to be representative of two different soil depths, hereafter called “surface” and “deep” soils, respectively (Fig. 1). For each of the five cores, we pooled portions of the same depth together to form one composite sample of ~200 g and we homogenized it. We took 15 g of soil from each composite sample and desiccated it immediately in sterile boxes with 40 g of silica gel. Previous analyses showed that this approach enables a cost-effective and long-term preservation of soil eDNA (Guerrieri et al., 2021). Soil eDNA collection was performed wearing gloves and the sampling tools were decontaminated with a portable blow torch (>1000°C) before the collection of each sample. We did not include soil litter and avoided roots, leaves and other large plant organs.

Soils were sampled “by depth” rather than “by horizons”, as is common practice in eDNA-based studies (Dickie et al., 2018) and in soil monitoring networks involving multiple glacier forelands (e.g., Khedim et al., 2021; Orgiazzi, Ballabio, Panagos, Jones, & Fernández-Ugalde, 2018; Rime et al., 2015; Schweizer, Hoeschen, Schlüter, Kögel-Knabner, & Mueller, 2018) because soil horizons are not yet differentiated in early stages of soil development, and because this approach allows obtaining a standardized pattern that can be applied across soils from multiple areas at very different development stages (Dickie et al., 2018; Khedim et al., 2021; Rime et al., 2015). Thus, the two categories “surface” (0-5 cm) and “deep” (7.5-20 cm) are used to define soil samples collected at two different soil depths, regardless of the horizons.

### Molecular analyses

In a dedicated room, we mixed the 15 g of soil with 20 ml of phosphate buffer for 15 min as described in Taberlet, Coissac, Pompanon, Brochmann, & Willerslev (2012); then we extracted eDNA using the Nucleo-Spin® Soil Mini Kit (Macherey-Nagel, Germany) with a final elution in 150 µl for both soil samples and with one negative extraction control every 23 samples (total: 12).

We amplified eDNA of bacteria, eukaryotes, fungi, springtails, insects and earthworms using primers designed for markers Bact02 (Bacteria and Archaea: Taberlet et al., 2018), Euka02 (Eukaryota: Guardiola et al., 2015), Fung02 (Mycota: Epp et al., 2012; Taberlet et al., 2018), Coll01 (Collembola, i.e. springtails: Janssen et al., 2018), Inse01 (Insecta: Taberlet et al., 2018), and Olig01 (Oligochaeta, i.e. earthworms: Bienert et al., 2012; Taberlet et al., 2018). We selected this set of markers to cover a wide range of organisms at different taxonomic resolution as we included three generalist markers (targeting entire superkingdoms or kingdoms: Bact02, Euka02 and Fung02) and three more specific markers (targeting from classes to

subclasses: Coll01, Inse01, Olig01). All these markers are well suited for metabarcoding analyses thanks to the low number of mismatches in the priming regions across target organisms, and they perform well with potentially degraded DNA due to the relatively short length of amplified fragments (Taberlet et al., 2018; Table S1). We used forward and reverse primers tagged on the 5'-end with eight-nucleotide long tags with at least five nucleotide differences among them (Coissac, 2012) and combined them in a way that all PCR replicates were represented by a unique combination of forward and reverse tags. This allowed us to uniquely identify each PCR replicate after sequencing. We randomized all samples on 96-well plates and included 24 bioinformatic blanks, 12 PCR negative controls and one PCR positive control. The positive control was constituted of genomic DNA of eight bacterial and two fungal strains (i.e., *Pseudomonas aeruginosa*, *Escherichia coli*, *Salmonella enterica*, *Lactobacillus fermentum*, *Enterococcus faecalis*, *Staphylococcus aureus*, *Listeria monocytogenes*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Cryptococcus neoformans*) at known concentrations (ZymoBIOMICS Microbial Community DNA Standard II, Zymo Research, USA; 1:10 diluted) and we used it to check for potential cross-contaminations and to monitor amplification and sequencing performances.

We determined the optimal number of amplification cycles and DNA dilution by conducting a qPCR essay on 48 randomly selected samples, using 1  $\mu$ l of 1:1000 diluted SYBR<sup>®</sup> Green I nucleic acid gel stain (Invitrogen, USA), and both undiluted and 1:10 diluted DNA, with a real-time PCR thermal cycler set to standard mode. This step is useful to avoid over-amplifying eDNA and to limit chimera formation.

Based on qPCRs results, we finally performed 42 (Bact02), 45 (Euka02, Fung02) or 55 (Coll01, Inse01, Olig01) amplification cycles on diluted (Euka02, Coll01) or undiluted (Bact02, Fung02, Inse01, Olig01) DNA. Amplification consisted of 20- $\mu$ l reactions with 10  $\mu$ l of AmpliTaq Gold 360 Master Mix 2X (Applied Biosystems, Foster City, CA, USA), 2  $\mu$ l of forward and reverse primer mix (initial concentration of each primer: 5  $\mu$ M), 0.16  $\mu$ l of Bovine Serum Albumin (i.e. 3.2  $\mu$ g; Roche Diagnostic, Basel, Switzerland) and 2  $\mu$ l of eDNA. We performed reactions in 384-well plates, with four PCR replicates per sample (Gentile F. Ficitola et al., 2015), setting the following PCR profiles: an initial step of 10 min at 95°C; several cycles of a 30 s denaturation at 95°C; a 30 s annealing at 53°C (Bact02), 45°C (Euka02), 56°C (Fung02), 51°C (Coll01), 55°C (Olig01) or 52°C (Inse01); a 90 s elongation for Bact02 and Fung02, or a 60 s elongation for all the others markers at 72°C; a final elongation at 72°C for 7 min. After amplification, we pooled together all amplicons of the same marker and visualized a 5- $\mu$ l aliquot by high-resolution capillary electrophoresis (QIAXcel Advanced System, QIAGEN, GERMANY) in order to check fragments length and monitor primer dimers. Finally, for each marker, we purified six subsamples of the pooled amplicons separately, using the MinElute PCR Purification Kit (QIAGEN, GERMANY) as per the manufacturer's instructions and combined them again. Libraries were prepared following the MetaFast protocol (Taberlet et al., 2018) and sequenced using the MiSeq (Bact02 and Fung02) or HiSeq 2500 (all others) Illumina platforms (Illumina, San Diego, CA, USA) with a paired-end approach (2  $\times$  250 bp for Bact02 and Fung02, and 2  $\times$  150 bp for the others markers) at Fasteris (SA, Geneva, Switzerland).

### Bioinformatic treatment

We used the OBITools software suite (Boyer et al., 2016) to perform the bioinformatic treatment of raw sequence data, as follows. First, we assembled the forward and reverse reads using the *illuminapairedend* program and kept only sequences with an alignment score greater than 40 (corresponding to a 10-nucleotide overlap of the forward and reverse reads). Second, we assigned aligned sequences to the corresponding PCR replicate using the program *ngsfilter* and allowed two and zero mismatches on primers and tags, respectively. Third, we dereplicated sequences using *obiuniq* and discarded bad-quality sequences (i.e., containing "N"), sequences whose length was lower or higher than expected (based on the minimum and maximum metabarcode length; Table S1) and singletons. Fourth, we ran the *obiclean* program with the option -r set at 0.5 to detect potential PCR or sequencing errors and kept only the sequences tagged as "heads" in at least one PCR. Sequences are tagged as "heads" when they are at least twice (-r option set at 0.5) as abundant as other related sequences differing by one base in the same PCR. Fifth, we clustered sequences at a threshold of 96% (Bact02, Euka02, Inse01), 95% (Fung02), 92% (Olig01) or 85% (Coll01) sequence similarity using

the *sumacrust* program (<https://git.metabarcoding.org/obitools/sumacrust/wikis/home>). These thresholds minimize the risk that sequences attributed to the same species are clustered in different MOTUs and were selected on the basis of preliminary bioinformatics analyses (Bonin, Guerrieri, & Ficetola, 2021).

For the taxonomic assignment, we built for each marker a sequence reference database from EMBL (version 140), as follows. First, we ran the *ecoPCR* program (Ficetola et al., 2010) to carry out *in silico* PCR with the primer pairs used for the experiment, allowing three mismatches per primer. Then, we curated the obtained reference databases by keeping only sequences assigned at the species, genus and family levels. Finally, the taxonomic assignment was performed by the *ecotag* program on each sequence using the reference database.

In order to remove spurious sequences and avoid bias in ecological conclusions (Calderón-Sanou, Munkemuller, Boyer, Zinger, & Thuiller, 2020) we performed additional filtering in R (version 4.0). We discarded MOTUs with a best identity < 80% (Bact02, Euka02, Fung02) or < 60% (Coll02, Inse01, Olig01) and MOTUs observed less than five (Bact02, Fung02, Inse01), ten (Olig01), eleven (Coll01) or twelve (Euka02) times overall. The latter corresponds to the minimum number of reads that removed [?] 99.99% of sequences detected in our blanks (i.e., tag-jump errors). Furthermore, we discarded MOTUs detected in only one sample, as they represent singletons, MOTUs detected in less than two PCR replicates of the same sample, as they often represent false positives (Gentile F. Ficetola et al., 2015), and MOTUs detected in more than one extraction or PCR negative control, as they might represent contaminants (Zinger, Bonin, et al., 2019).

### Statistical analyses

At each plot and for each depth, we measured alpha-diversity of Bacteria, Eukaryota, Mycota, Collembola, Insecta and Oligochaeta, through Hill numbers. The joint use of different Hill numbers allows to obtain biodiversity measures in metabarcoding studies that are robust to bioinformatic treatments and other methodological choices (Alberdi & Gilbert, 2019; Calderon-Sanou et al., 2020; Machler, Walser, & Altermatt, 2021). We used the parameters  $q = 0$  and  $q = 1$  in the *hill\_taxa* function of the *hillR* package (Chao, Chiu, & Jost, 2014). Values of  $q = 0$  returns the taxonomic richness and are insensitive to MOTUs frequency, while  $q = 1$  returns the exponential of the Shannon diversity, and limits the weight of rare MOTUs (Alberdi & Gilbert, 2019). We could not use values of  $q > 1$  because they cannot be applied to communities with richness = 0.

We used univariate Bayesian Generalized Linear Mixed Models (GLMMs) to assess the variation in alpha-diversity of Bacteria, Eukaryota, Mycota, Collembola, Insecta and Oligochaeta with time since glacier retreat and depth. We ran mixed models with the alpha-diversity of each sample (log-transformed) as dependent variable, and used a Gaussian error to attain normality of model residuals, considering both the parameters  $q = 0$  and  $q = 1$ . As independent variables, we considered time since glacier retreat (log-transformed and scaled: mean = 0, SD = 1) and depth. We included glacier identity and plot identity as random factors. These models also included interactions between depth and time since glacier retreat, to test the hypothesis that depth affects the colonization rate of the studied groups. We used the widely applicable information criterion (WAIC) to compare models with and without interactions (Gelman, Hwang, & Vehtari, 2013). Models using log-normal error distribution and un-transformed alpha diversity values yielded identical results.

Soil nutrient content changes at different stages of soil development in deglaciated areas, with the amount of organic carbon increasing through time (Khedim et al., 2021), potentially influencing alpha diversity of soil communities (Guo et al., 2018). We therefore re-analysed the pattern of alpha diversity using organic carbon as an independent variable, instead of age since glacier retreat. This analysis was run for a subset of samples ( $N = 276$ ) for which data of total organic carbon content were obtained by elemental analysis (Organic Elemental Analyzer, model Flash 2000, Thermo Fisher; Khedim et al., 2021; Lacchini 2020). Organic carbon was strongly related to age since glacier retreat (GLMM with organic carbon as dependent variable and age as independent variable:  $R^2 = 0.66$ ) thus it was impossible to include organic carbon and age together in the same model. Organic carbon data were representative of the overall soil core (0-20 cm); thus, the analysis did not allow testing the role of variation in carbon content between surface and deep layers. Two plots (i.e. four samples in total) were excluded from this analysis because no soil data were

available.

For each plot, we estimated the beta-diversity between the two soil depths based on incidence data. We used the *beta.multi* function of the *betapart* package with the Sorensen family (Baselga & Orme, 2012). This function partitions the total beta diversity (beta.SOR) into its nestedness (beta.SNE) and turnover (beta.SIM) components, reflecting the species gain/loss and replacement, respectively (Baselga, 2010). We excluded plots having zero MOTUs in at least one depth, given that the formula of Baselga’s partitioning retrieves undefined values of nestedness and turnover when one of the compared communities has no taxon (Baselga, 2010). For each taxonomic group, we built GLMMs to test the hypothesis that beta diversity between the two soil depths decreases with time since glacier retreat. We ran mixed models with rescaled indices (Smithson & Verkuilen, 2006) to avoid fixed zeros and ones, using a beta distribution, and included glacier identity as a random factor. Models for beta diversity were limited to plots with at least one detected MOTU in both depths. We then repeated the analyses for the turnover and nestedness components of beta diversity. We ran all generalized mixed models with three MCMC chains, 5,000 iterations and a burn-in of 5,000 in the *brms* R package (Burkner, 2017). After processing, c-hat values were always <1.01, indicating convergence.

To visualize the variation of the structure of belowground communities across different stages of soil development, we used distance-based Principal Component Analysis (db-PCA). We calculated distance among samples using the Hellinger distance that allows us to control for the double zero issue (Legendre & Legendre, 2012). Prior to ordination, count data were normalized with a shift-log transformation in order to stabilize extreme values and variance inflation. As for the beta-diversity analysis, we removed plots having zero MOTUs in at least one depth. To test for differences in communities across time, depth and their potential interaction, we performed a permutational multivariate analysis of variance (PERMANOVA) using the *adonis* function of the *vegan* package (Oksanen et al., 2019) with glaciers as strata and permutations set to 9999. Time was log-transformed. Results of PERMANOVA can be sensitive to differences in multivariate dispersion (Anderson, 2001), therefore we computed the homogeneity of variance among groups (Anderson, 2006) and tested for its significance by permutations ( $n = 9999$ ). We used data visualization in ordination plots to support the interpretation of the statistical tests.

Finally, we used the indicator value (IndVal; Dufrene & Legendre 1997) approach to identify MOTUs that were characteristic for particular stages of soil development and/or soil depth. Prior to the analysis, we grouped samples into three classes based on the years since glacier retreat (i.e., < 40 years; 40-95 years and > 95 years) and depth (surface/deep), for a total of six environmental classes. Metabarcoding approaches can lead to a very large number of MOTUs. Thus, for this analysis we only retained MOTUs with a relative abundance > 0.1% for each taxonomic group. We computed the IndVal index using the *indicspecies* package (De Caceres & Legendre, 2009). For a given taxon, the IndVal index is based on its specificity (i.e., the concentration of abundance) and its fidelity (i.e., the relative frequency) within a class. Each MOTU could be an indicator of at maximum two environmental classes (De Caceres, Legendre, & Moretti, 2010), so that a MOTU could result as indicator of e.g. one or both depths at a given soil stage, or of one or two consecutive stages at a given soil depth. This choice allowed keeping the number and the ecological meaningfulness of the combinations reasonable. The significance of indicator values was tested through random permutations ( $n = 9999$ ) and  $p$ -values were adjusted for multiple comparison tests using the FDR method (Benjamini & Hochberg, 1995). We used the packages *ggplot2* (Wickham, 2016), *ggpubr* (Kassambara, 2020), *phyloseq* (McMurdie & Holmes, 2013) and *vegan* (Oksanen et al., 2019) for multivariate statistical analyses and visualization.

## RESULTS

A total of 7,335,969 (Bact02), 9,655,151 (Euka02), 15,649,401 (Fung02), 8,941,690 (Coll01), 7,537,839 (Inse01) and 6,107,034 (Olig01) reads were obtained after bioinformatic filtering. After clustering and spurious sequence removal, DNA metabarcoding yielded 1,825 (Bact02), 753 (Euka02), 1,483 (Fung02), 118 (Coll01), 396 (Inse01) and 97 (Olig01) MOTUs.

## How is alpha-diversity related to soil depth, time since deglaciation and soil features?

Overall, the alpha-diversity was highest for the generalist markers (Euka02, Bact02, Fung02) compared to the specialist ones (Coll01, Inse01, Olig01). Estimates of alpha-diversity obtained with different Hill numbers ( $q = 0$  and  $q = 1$ ) were strongly correlated (for all taxonomic groups,  $r > 0.78$ ; Table S2).

When we used  $q = 1$ , we observed an increase of alpha-diversity with time since glacier retreat for all the taxonomic groups. For Eukaryota and Mycota, alpha-diversity was significantly higher in communities retrieved at surface layers with depth of 0-5 cm, compared to the communities detected in the deeper layer of soil (Fig. 2; Table S3). Furthermore, for Collembola we detected an interaction between depth and time since glacier retreat. For this group, the alpha-diversity index was close to one (mean:  $1.19 \pm 0.51$ ; corresponding to richness  $\sim 0$ ) at relatively young sites ( $< 30$  years after glacier retreat) and increased with time, but the increase was faster in communities at 0-5 cm of depth. All results were highly consistent when we repeated analyses using  $q = 0$  (Table S3). Results were very similar when we used soil carbon content as a predictor variable instead of time since glacier retreat, as GLMMs showed a significant increase in alpha-diversity with average organic carbon content of the plot, even though the  $R^2$  values of these models were generally lower than the  $R^2$  of models with age as independent variable (Fig. S1; Table S4).

## Changes in beta diversity through time

In order to assess the beta diversity between surface and deep soils, plots having zero MOTUs in at least one depth were removed, corresponding to 4.28% (Bacteria), 14.28% (Eukaryota), 6.43% (Mycota), 39.28% (Collembola), 13.57% (Insecta) and 37.85% (Oligochaeta) of total plots.

GLMMs allowed us to detect changes in the beta-diversity of communities between surface and deep soil. Differences in community composition between the two depths decreased with time since glacier retreat for Bacteria, Eukaryota, Mycota and Insecta, indicating homogenization of communities, while we did not detect significant changes through time for the beta-diversity of Collembola and Oligochaeta (Fig. 3; Supplementary Table S5). Collembola and Oligochaeta were also the taxa for which the largest number of sites were discarded because of a lack of MOTUs. Overall, our models did not show significant changes in the turnover or nestedness components of the beta diversity measures through time, with the only exception of Oligochaeta, for which nestedness between surface and deep soils tended to increase through time (Supplementary Table S5 and Fig. S2).

Within each deglaciated foreland, the structure of communities was primarily related to time since glacier retreat (Fig. 4). Time significantly affected community structure for Bacteria, Mycota and Eukaryota (PERMANOVA:  $p < 0.05$ ; Table 1); the amount of variance explained by time ranged from 2.4% to 5.7%. For Bacteria, Mycota, Eukaryota, as well as for Insecta, community structure also differed significantly between soil depths, but the explained variance was smaller ( $< 1\%$ ; Table 1). For none of the groups, we detected a significant interaction between time and soil depth (Table 1), suggesting that the effect of time was consistent between surface and deep soils. Differences in multivariate dispersions were never significant between soil depths, but were significant across time except for Collembola (Table 1). Bacterial community structure was the most strongly related to time and depth ( $R^2 = 5\%$ ; Table 1). Differences among deglaciated forelands were marked but tended to follow similar trends across the taxonomic groups (Fig. 4).

Based on the specificity and fidelity of each MOTU, 86 were identified as indicators (47 Bacteria, 34 Mycota and five Eukaryota; Table S6). For Bacteria, 22 taxa were strongly associated with young foreland soils, including members of the genera *Roseiflexus*, *Herbaspirillum*, *Novosphingobium* that exhibited particularly high IndVal, while no one was strictly associated to the intermediate ages. Seventeen taxa of Bacteria were indicators of both surface and deep soil layers in older forelands, including members from the genera *Actinoallomurus* and *Ferrimicrobium* that showed the highest IndVal. Six taxa were indicators of the deep soil layers at both intermediate and old age. For Eukaryota, five taxa were considered as indicators; three were fungi related to old soils, while one mite (genus *Gamasina*) was associated with the intermediate age class. For Mycota, 18 taxa were indicators of both surface and deep layers in older forelands, including members of the genus *Cladophialophora* and the family *Glomeraceae*. Ten Mycota taxa were indicators of

both surface and deep layers in young forelands while intermediate forelands contained less indicators, with only five taxa. Only one Mycota taxon was representative of a specific soil layer (the MOTU identified as *Golovinomyces sordidus*, associated to the surface layer of young forelands).

## DISCUSSION

Our work provides new insights on the colonization and primary succession patterns in deglaciated terrains, by integrating soil depth in the primary succession studies and by implementing a multi-taxa approach across multiple forelands. Here, the multi-taxa approach allowed characterizing patterns for a wide range of topsoil organisms involved in colonization and successional processes. In order to cover the largest proportion of biodiversity, we considered both generalist (for Bacteria, Mycota and Eukaryota) and more specific (for Collembola, Insecta and Oligochaeta) markers. Alpha- and beta-diversity variation through time showed a strikingly consistent pattern across these taxa. The considered depths did not strongly affect the alpha diversity of some taxa at any stage of soil development, even though communities inhabiting surface and deep soil layers were not exactly the same. Importantly, beta-diversity between surface and deep soil decreased through time across most of taxa, supporting the hypothesis of homogenization between surface and deep soil along the succession (Rime et al., 2015).

### Changes in alpha-diversity with soil age and the impact of depth

Alpha-diversity increased through time, as previously observed in successional studies of microorganisms, plants and soil invertebrates (Erschbamer & Caccianiga, 2016; Ficetola et al., 2021; Matthews, 1992), with a similar pattern across all the study groups. For the whole Eukaryota and, within them, for Mycota, the highest alpha-diversity was found in the surface soils, supporting our hypothesis that the richness of communities decreases toward deep soils. This observation agrees with the idea that the highest soil biodiversity is hosted close to the surface, as already observed for fungi, bacteria and some faunal groups (Carteron et al., 2021; Chen et al., 2020; Chu et al., 2016; Jiao et al., 2018; Moradi et al., 2020; Mundra et al., 2021; Rime et al., 2015). In glacier forelands, soils tend to have higher water holding capacity, more exchangeable cations, carbon and nutrient contents toward the surface (Rime et al., 2015). These properties are vital for most belowground organisms, especially in those resource-limited ecosystems, resulting in higher bacterial activity, DNA concentration, fungal and root biomass in the first centimeters of soil (Rime et al., 2015). We highlight that, in our sampling design, the surface sampling covered a thinner layer compared to the deep one (from 0 to 5 cm vs. from 7.5 to 20 cm of depth). In principle, the deep layer might hold larger environmental heterogeneity, given that it is the thickest one. Thus, alpha-diversity between layers might be even larger, had we sampled layers with similar thicknesses.

For springtails only, the interaction between soil depth and development stage had a significant effect on alpha-diversity, indicating that for this group taxonomic richness increased at different rates between the two soil depths. Springtails were nearly absent in soils aged less than 30 years (Fig. 2). Then, the alpha-diversity increased, but the increase was faster in the surface layer compared to the deep layer, probably because the fast accumulation of organic matter in surface soils (Herold et al., 2014; Moradi et al., 2020) allows the establishment of these organisms, which have multiple trophic roles, from detritivore to herbivorous. For the other taxa, we did not detect significant interaction between soil depth and soil development stage, suggesting that alpha-diversity increases through time with a similar pattern between surface and deep layers, except for Collembola and perhaps in very young soils (see Fig. 2).

In glacier forelands, the amount of organic matter consistently increases through time (Khedim et al., 2021). By repeating the analyses of alpha-diversity, considering the organic carbon content as independent variable instead of time since glacier retreat, we confirmed that our conclusions are not biased by the issues of using different sites as substitutes of time (issues of space-for-time substitution in successional studies; Johnson & Miyanishi, 2008). Soil carbon content is a major driver of soil biodiversity changes (Chu et al., 2016); consistently with this idea, alpha-diversity tended to increase with organic carbon. Nevertheless, models with time showed slightly higher  $R^2$  values than the ones with soil organic carbon (Fig. 2, supplementary Figure S1), suggesting that time since glacier retreat is a better predictor of alpha-diversity than organic carbon,

even though these parameters are strongly correlated (Rime et al., 2015; Zumsteg et al., 2011). Further studies are needed in order to disentangle the role of both time and soil features as drivers of primary succession.

### **Communities differences between surface and deep soils change through time**

The beta-diversity between surface and deep layers was particularly high soon after the retreat of glaciers, and then decreased with time. As seen for the alpha diversity, this pattern was consistent across nearly all taxa. Collembola and Oligochaeta are the only taxa for which this was not evident, but these animals were nearly absent from recently deglaciated soils (and particularly from the deep layers; Fig. 2), probably because many of them require well developed soils, with abundant organic matter to find resources (Phillips et al., 2019). Therefore, for Collembola and Oligochaeta, many plots at early development stages were excluded from this analysis, reducing statistical power. In principle, the variation of beta diversity between surface and deep layer can be attributable to both species gain/loss (nestedness) and replacement (Baselga, 2010). Turnover was more important than nestedness for invertebrates (Collembola, Insecta, Oligochaeta), while for microorganisms (Bacteria and Mycota) turnover and nestedness showed a similar importance (Supplementary fig. S2), and the relevance of these two components of beta-diversity remained similar through time (Supplementary Fig. S2; Table S5).

The decrease in beta-diversity between surface and deep layers through time confirms the hypothesis of homogenization of communities (Rime et al., 2015), and extends it to the whole soil biota, as bacteria, microeukaryotes and animals responded the same way (Fig. 3). Community homogenization is probably related to the structural modifications observed during the development of soil horizons (e.g. Schaetzl & Anderson, 2005). The study of sites at different stages of soil formation has shown a differentiation of organic horizon immediately after glacial retreat (O), followed by the development of an organo-mineral horizon (A) during the first 150 years (Crocker & Major, 1955; Mavris, Egli, Plotze, Blum, Mirabella and Giaccari, 2010). The strong vertical variation of physical, chemical and structural features (e.g. light, temperature, pH; Moradi et al., 2020; Mundra et al., 2021) clearly affects communities, which show a particularly strong response to fine-scale environmental heterogeneity (Rime et al., 2015; Moradi et al., 2020; Mundra et al., 2021). For example, immediately after glacier retreat, the amount of fine sediments is the highest at the surface (Rime et al., 2015). This can determine differences in humidity between the surface and the deeper layers, that in turn affect communities (Rime et al., 2015). The decrease of beta-diversity can be explained by the progressive deepening of the organo-mineral horizon (Mavris et al., 2010), where abundant resources favor the establishment of complex communities. Plant richness and cover quickly increase during the first decades after glacier retreat, and 40 years after glacier retreat plants cover generally rises above 50% (Rime et al., 2015). Plant roots generally influence the first 20 cm of soils and more, and could have determined the homogenization of superficial and deep samples of our study (Rime et al., 2015). Differences between surface and deep layers would probably be stronger if a larger vertical gradient is analyzed (e.g. from surface to 50 cm deep; Moradi et al., 2020), and this is certainly an important aspect that deserves future studies. However, in glacier forelands the study of deep layers by eDNA analysis is sometimes problematic because rock outcrops are frequent a few centimeters below the surface. In any case, for all the taxa considered, time since glacier retreat remained the main determinant of community variation, as it explained much more variation in community composition compared to depth (Table 1). This confirms the idea that, even though fine-scale heterogeneity certainly has a role, time since glacier retreat remains the main determinant of community evolution after glacier retreat (Ficetola et al., 2021; Rime et al., 2015).

For microorganisms, the significant community differences between soil layers (Table 1) likely are determined by taxa that are specialists of given environmental features. This idea is supported by the observation that all MOTUs identified as indicators of surface or deep soils are bacteria or fungi (supplementary Table S6). Conversely, for invertebrates, soil depth explained a very limited amount of variation in community composition (Table 1). This could be due to the lower richness of these taxa (which limits statistical power), or to the fact that a broader vertical profile would be required to identify specialists (Moradi et al., 2020). Several taxa identified as indicators in Rime et al. (2015) showed similar patterns across the

different locations of our study, confirming the strong functional variation of communities through time. Taxa identified as clear indicators both here and by Rime et al. (2015) include the *Clostridium* bacteria, known as anaerobic, which were consistently found as indicators of the earliest stage of soil development. Similarly, several fungal saprotrophs were indicators across the different stages of soil development, while *Lachnum* was a microfungus consistently associated with the most developed soils (Nguyen et al., 2016). *Gemmatimonas* tend to be copiotrophic bacteria (Ho, Di Lonardo, & Bodelier, 2017) and include multiple MOTUs that were found as indicators of different stages. Interestingly, fungi such as *Laccaria* or *Hygrophorus* (i.e., potential ectomycorrhizal taxa; Nguyen et al., 2016) were also indicators of later stages of development. Contrary to Rime et al. (2015), arbuscular mycorrhizal fungi (i.e., Glomeromycetes) tended to be associated with oldest forelands, confirming the growing importance of plant-associated fungi along community development (Davey et al., 2015).

## Conclusion

Understanding the development of communities in primary successions remains a major task of ecological studies. Our study suggests that, even though time since glacier retreat is a more important driver than depth in shaping the diversity of communities, patterns are not identical for superficial and deeper samples. This can have important consequences on ecosystem functioning, for example for the sequestration of organic carbon in these soils (Khedim et al., 2021). Nevertheless, differences between depths tend to decrease through time with a consistent pattern in both microorganisms and animals, possibly because of the increasing role of plants along successional stages. Future studies are required to identify possible factors driving biotic colonization within the same system (e.g. microclimate, soil features, etc.) and patterns of biotic interactions.

## ACKNOWLEDGMENTS

This study is funded by the European Research Council under the European Community's Horizon 2020 Programme, Grant Agreement no. 772284 (IceCommunities).

## REFERENCES

- Alberdi, A., & Gilbert, M. T. P. (2019). A guide to the application of Hill numbers to DNA-based diversity analyses. *Molecular Ecology Resources* , 19 (4), 804–817. doi: 10.1111/1755-0998.13014
- Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology* , 26 (1), 32–46. doi: 10.1111/J.1442-9993.2001.01070.PP.X
- Anderson, M. J. (2006). Distance-Based Tests for Homogeneity of Multivariate Dispersions. *Biometrics* , 62 (1), 245–253. doi: 10.1111/J.1541-0420.2005.00440.X
- Bajerski, F., & Wagner, D. (2013). Bacterial succession in Antarctic soils of two glacier forefields on Larsemann Hills, East Antarctica. *FEMS Microbiology Ecology* , 85 (1), 128–142. doi: 10.1111/1574-6941.12105
- Bardgett, R. D., & Van Der Putten, W. H. (2014). Belowground biodiversity and ecosystem functioning. *Nature* , 515 . doi: 10.1038/nature13855
- Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* , 19 (1), 134–143. doi: 10.1111/j.1466-8238.2009.00490.x
- Baselga, A., & Orme, C. D. L. (2012). Betapart: An R package for the study of beta diversity. *Methods in Ecology and Evolution* , 3 (5), 808–812. doi: 10.1111/j.2041-210X.2012.00224.x
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* , 57 (1), 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J.-J., & Taberlet, P. (2012). Tracking earthworm communities from soil DNA. *Molecular Ecology* , 21 (8), 2017–2030. doi: 10.1111/j.1365-294X.2011.05407.x

- Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., . . . de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution* . doi: 10.1016/j.tree.2014.04.003
- Bonin, A., Guerrieri, A., & Ficetola, G. F. (2021). Optimal sequence similarity thresholds for clustering of molecular operational taxonomic units in DNA metabarcoding studies. *Authorea Preprints* . doi: 10.22541/AU.163398436.69595421/V1
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources* , 16 (1), 176–182. doi: 10.1111/1755-0998.12428
- Burkner, P. C. (2017). Advanced Bayesian Multilevel Modeling with the R Package brms. *R Journal* , 10 (1), 395–411. doi: 10.32614/rj-2018-017
- Calderon-Sanou, I., Munkemuller, T., Boyer, F., Zinger, L., & Thuiller, W. (2020). From environmental DNA sequences to ecological conclusions: How strong is the influence of methodological choices? *Journal of Biogeography* , 47 (1), 193–206. doi: 10.1111/jbi.13681
- Carteron, A., Beigas, M., Joly, S., Turner, B. L., & Laliberte, E. (2021). Temperate Forests Dominated by Arbuscular or Ectomycorrhizal Fungi Are Characterized by Strong Shifts from Saprotrophic to Mycorrhizal Fungi with Increasing Soil Depth. *Microbial Ecology* ,82 (2), 377–390. doi: 10.1007/s00248-020-01540-7
- Cauvy-Fraunie, S., & Dangles, O. (2019). A global synthesis of biodiversity responses to glacier retreat. *Nature Ecology and Evolution* , 3 (12), 1675–1685. doi: 10.1038/s41559-019-1042-8
- Chao, A., Chiu, C.-H., & Jost, L. (2014). Unifying Species Diversity, Phylogenetic Diversity, Functional Diversity, and Related Similarity and Differentiation Measures Through Hill Numbers. *Http://Dx.Doi.Org/10.1146/Annurev-Ecolsys-120213-091540* ,45 , 297–324. doi: 10.1146/ANNUREV-ECOLSYS-120213-091540
- Chen, W., Jiao, S., Li, Q., & Du, N. (2020). Dispersal limitation relative to environmental filtering governs the vertical small-scale assembly of soil microbiomes during restoration. *Journal of Applied Ecology* , 57 (2), 402–412. doi: 10.1111/1365-2664.13533
- Chu, H., Sun, H., Tripathi, B. M., Adams, J. M., Huang, R., Zhang, Y., & Shi, Y. (2016). Bacterial community dissimilarity between the surface and subsurface soils equals horizontal differences over several kilometers in the western Tibetan Plateau. *Environmental Microbiology* , 18 (5), 1523–1533. doi: 10.1111/1462-2920.13236
- Coissac, E. (2012). OligoTag: A program for designing sets of tags for next-generation sequencing of multiplexed samples. *Methods in Molecular Biology* , 888 , 13–31. doi: 10.1007/978-1-61779-870-2\_2
- Crocker, R. L., & Major, J. (1955). Soil Development in Relation to Vegetation and Surface Age at Glacier Bay, Alaska. *Journal of Ecology*, 43(2), 427–448. doi:10.2307/2257005
- Davey, M., Blaaid, R., Vik, U., Carlsen, T., Kauserud, H., & Eidesen, P. B. (2015). Primary succession of *Bistorta vivipara* (L.) Delabre (Polygonaceae) root-associated fungi mirrors plant succession in two glacial chronosequences. *Environmental Microbiology* ,17 (8), 2777–2790. doi: 10.1111/1462-2920.12770/SUPPINFO
- De Caceres, M., & Legendre, P. (2009). Associations between species and groups of sites: Indices and statistical inference. *Ecology* ,90 (12), 3566–3574. doi: 10.1890/08-1823.1
- De Caceres, M., Legendre, P., & Moretti, M. (2010). Improving indicator species analysis by combining groups of sites. *Oikos* ,119 (10), 1674–1684. doi: 10.1111/j.1600-0706.2010.18334.x
- Dickie, I. A., Boyer, S., Buckley, H. L., Duncan, R. P., Gardner, P. P., Hogg, I. D., . . . Weaver, L. (2018). Towards robust and repeatable sampling methods in eDNA-based studies. *Molecular Ecology Resources* , 18

(5), 940–952. doi: 10.1111/1755-0998.12907

Doblas-Miranda, E., Sanchez-Pinero, F., & Gonzalez-Megias, A. (2009). Vertical distribution of soil macrofauna in an arid ecosystem: Are litter and belowground compartmentalized habitats? *Pedobiologia* ,52 (6), 361–373. doi: 10.1016/J.PEDOBI.2008.11.006

Donald, J., Murienne, J., Chave, J., Iribar, A., Louisanna, E., Manzi, S., ... Zinger, L. (2021). Multi-taxa environmental DNA inventories reveal distinct taxonomic and functional diversity in urban tropical forest fragments. *Global Ecology and Conservation* , 29 , e01724. doi: 10.1016/J.GECCO.2021.E01724

Dufrene, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, 67, 345–366. doi:/10.2307/2963459

Erschbamer, B., Caccianiga, M., 2016. Glacier Forelands: Lessons of Plant Population and Community Development. *Progress in Botany* 78, 259-284. doi: 10.1007/124\_2016\_4

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., ... Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Molecular Ecology* , 21 (8), 1821–1833. doi: 10.1111/j.1365-294X.2012.05537.x

Ficetola, Gentile F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessiere, J., ... Pompanon, F. (2010). An In silico approach for the evaluation of DNA barcodes. *BMC Genomics* , 11 (1), 434. doi: 10.1186/1471-2164-11-434

Ficetola, Gentile F., Pansu, J., Bonin, A., Coissac, E., Giguët-Covex, C., De Barba, M., ... Taberlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources* . doi: 10.1111/1755-0998.12338

Ficetola, Gentile F., Marta, S., Guerrieri, A., Gobbi, M., Ambrosini, R., Fontaneto, D., ... Thuiller, W. (2021). Dynamics of Ecological Communities Following Current Retreat of Glaciers. *Annual Review of Ecology, Evolution, and Systematics* , 52 (1), 1–51. doi: 10.1146/annurev-ecolsys-010521-040017

Fierer, N., Schimel, J. P., & Holden, P. A. (2003). Variations in microbial community composition through two soil depth profiles. *Soil Biology and Biochemistry* , 35 (1), 167–176. doi: 10.1016/S0038-0717(02)00251-1

Franzetti A, Pittino F, Gandolfi I, Azzoni RS, Diolaiuti G, Smiraglia C, Pelfini M, Compostella C, Turchetti B, Buzzini P, Ambrosini R. (2020). Early ecological succession patterns of bacterial, fungal and plant communities along a chronosequence in a recently deglaciated area of the Italian Alps. *FEMS Microbiol Ecol* , 96(10):fiaa165. doi: 10.1093/femsec/fiaa165. PMID: 32815995.

Gelman, A., Hwang, J., & Vehtari, A. (2013). Understanding predictive information criteria for Bayesian models. *Statistics and Computing* 2013 24:6 , 24 (6), 997–1016. doi: 10.1007/S11222-013-9416-2

Gobbi, M., Ballarin, F., Brambilla, M., Compostella, C., Isaia, M., Losapio, G., ... Caccianiga, M. (2017). Life in harsh environments: carabid and spider trait types and functional diversity on a debris-covered glacier and along its foreland. *Ecological Entomology* , 42 (6), 838–848. doi: 10.1111/EEN.12456

Guardiola, M., Uriz, M. J., Taberlet, P., Coissac, E., Wangensteen, O. S., & Turon, X. (2015). Deep-Sea, Deep-Sequencing: Metabarcoding Extracellular DNA from Sediments of Marine Canyons. *PLOS ONE* ,10 (10), e0139633. doi: 10.1371/journal.pone.0139633

Guerrieri, A., Bonin, A., Munkemüller, T., Gielly, L., Thuiller, W., & Francesco Ficetola, G. (2021). Effects of soil preservation for biodiversity monitoring using environmental DNA. *Molecular Ecology* . doi: 10.1111/mec.15674

Guo, Y., Chen, X., Wu, Y., Zhang, L., Cheng, J., Wei, G., & Lin, Y. (2018). Natural revegetation of a semiarid habitat alters taxonomic and functional diversity of soil microbial communities. *Science of The Total Environment* , 635 , 598–606. doi: 10.1016/J.SCITOTENV.2018.04.171

- Hagvar, S., Gobbi, M., Kaufmann, R., Ingimarsdottir, M., Caccianiga, M., Valle, B., ... Vater, A. (2020). *Ecosystem Birth Near Melting Glaciers: A Review on the Pioneer Role of Ground-Dwelling Arthropods* .11 , 644. doi: 10.3390/insects11090644
- Herold, N., Schoning, I., Berner, D., Haslwimmer, H., Kandeler, E., Michalzik, B., & Schrumpf, M. (2014). Vertical gradients of potential enzyme activities in soil profiles of European beech, Norway spruce and Scots pine dominated forest sites. *Pedobiologia* , 57 (3), 181–189. doi: 10.1016/J.PEDOBI.2014.03.003
- Ho, A., Di Lonardo, D. P., & Bodelier, P. L. E. (2017). Revisiting life strategy concepts in environmental microbial ecology. *FEMS Microbiology Ecology* , 93 (3). doi: 10.1093/FEMSEC/FIX006
- Ingham, R. E., Trofymow, J. A., Ingham, E. R., & Coleman, D. C. (1985). Interactions of Bacteria, Fungi, and their Nematode Grazers: Effects on Nutrient Cycling and Plant Growth. *Ecological Monographs* ,55 (1), 119–140. doi: 10.2307/1942528
- Janssen, P., Bec, S., Fuhr, M., Taberlet, P., Brun, J.-J., & Bouget, C. (2018). Present conditions may mediate the legacy effect of past land-use changes on species richness and composition of above- and below-ground assemblages. *Journal of Ecology* , 106 (1), 306–318. doi: 10.1111/1365-2745.12808
- Jiao, S., Chen, W., Wang, J., Du, N., Li, Q., & Wei, G. (2018). Soil microbiomes with distinct assemblies through vertical soil profiles drive the cycling of multiple nutrients in reforested ecosystems. *Microbiome* , 6 (1), 1–13. doi: 10.1186/s40168-018-0526-0
- Johnson, E. A., & Miyanishi, K. (2008). Testing the assumptions of chronosequences in succession. *Ecology Letters* , 11 (5), 419–431. doi: 10.1111/J.1461-0248.2008.01173.X
- Kassambara, A. (2020). “*ggplot2*” Based Publication Ready Plots [R package ggpubr version 0.4.0] . Retrieved from <https://cran.r-project.org/package=ggpubr>
- Kaufmann, R. (2001). Invertebrate succession on an alpine glacier foreland. *Ecology* , 82 (8), 2261–2278. doi: 10.1890/0012-9658(2001)082[2261:ISOAAG]2.0.CO;2
- Khedim, N., Cecillon, L., Poulenard, J., Barre, P., Baudin, F., Marta, S., ... Ficetola, G. F. (2021). Topsoil organic matter build-up in glacier forelands around the world. *Global Change Biology* ,27 (8), 1662–1677. doi: 10.1111/gcb.15496
- Khokon, A. M., Schneider, D., Daniel, R., & Polle, A. (2021). Soil Layers Matter: Vertical Stratification of Root-Associated Fungal Assemblages in Temperate Forests Reveals Differences in Habitat Colonization. *Microorganisms* 2021, Vol. 9, Page 2131 ,9 (10), 2131. doi: 10.3390/MICROORGANISMS9102131
- Lacchini, F. 2020. L’evoluzione del suolo nelle piane proglaciali in relazione al cambiamento climatico. Unpublished MSc Thesis, Università degli Studi di Milano
- Legendre, P., & Legendre, L. (2012). Numerical Ecology Ch 6 - Multidimensional qualitative data. *Developments in Environmental Modelling* , 24 , 337–424. doi: 10.1016/B978-0-444-53868-0.50008-3
- Machler, E., Walser, J.-C., & Altermatt, F. (2021). Decision-making and best practices for taxonomy-free environmental DNA metabarcoding in biomonitoring using Hill numbers. *Molecular Ecology* ,30 (13), 3326–3339. doi: 10.1111/MEC.15725
- Marta, S., Azzoni, R. S., Fugazza, D., Tielidze, L., Chand, P., Sieron, K., ... Zimmer, A. (2021). *The Retreat of Mountain Glaciers since the Little Ice Age : A Spatially Explicit Database* . 4–11.
- Matthews, J.A., 1992. The Ecology of Recently-deglaciated Terrain: A Geocological Approach to Glacier Forelands and Primary Succession. Cambridge University Press, Cambridge. doi.org/10.2307/2261511
- Mavris, C., Egli, M., Plotze, M., Blum, J.D., Mirabella, A., Giaccai, D., Haeberli, W. (2010): Initial stages of weathering and soil formation in the Morteratsch proglacial area (Upper Engadine, Switzerland). *Geoderma*, 155, 359–371. doi:10.1016/j.geoderma.2009.12.019

- McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* , 8 (4), e61217. doi: 10.1371/JOURNAL.PONE.0061217
- Menta, C. Soil fauna diversity-function, soil degradation, biological indices, soil restoration. In *Biodiversity Conservation and Utilization in a Diverse World*; Lameed, G.A., Ed.; IntechOpen: London, UK, 2012; pp. 59–94. doi: 10.5772/51091
- Moradi, J., John, K., Vicentini, F., Vesela, H., Vicena, J., Ardestani, M. M., & Frouz, J. (2020). Vertical distribution of soil fauna and microbial community under two contrasting post mining chronosequences: Sites reclaimed by alder plantation and unreclaimed regrowth. *Global Ecology and Conservation* , 23 . doi: 10.1016/j.gecco.2020.e01165
- Mundra, S., Kjonaas, O. J., Morgado, L. N., Krabberod, A. K., Ransedokken, Y., & Kauserud, H. (2021). Soil depth matters: shift in composition and inter-kingdom co-occurrence patterns of microorganisms in forest soils. *FEMS Microbiology Ecology* , 97 (3). doi: 10.1093/FEMSEC/FIAB022
- Nguyen, N. H., Song, Z., Bates, S. T., Branco, S., Tedersoo, L., Menke, J., ... Kennedy, P. G. (2016). FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecology* , 20 , 241–248. doi: 10.1016/J.FUNECO.2015.06.006
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., ... Maintainer, H. W. (2019). *Package “vegan” Title Community Ecology Package Version 2.5-6* .
- Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., & Fernandez-Ugalde, O. (2018). LUCAS Soil, the largest expandable soil dataset for Europe: a review. *European Journal of Soil Science* , 69 (1), 140–153. doi: 10.1111/EJSS.12499
- Orwin, K. H., Kirschbaum, M. U. F., John, M. G. S., & Dickie, I. A. (2011). Organic nutrient uptake by mycorrhizal fungi enhances ecosystem carbon storage: a model-based assessment. *Ecology Letters* , 14 (5), 493–502. doi: 10.1111/J.1461-0248.2011.01611.X
- Phillips, H. R. P., Guerra, C. A., Bartz, M. L. C., Briones, M. J. I., Brown, G., Crowther, T. W., ... Eisenhauer, N. (2019). Global distribution of earthworm diversity. *Science* , 366 (6464), 480–485. doi: 10.1126/SCIENCE.AAX4851/SUPPL\_FILE/AAX4851-PHILLIPS-SM.PDF
- Rime, T., Hartmann, M., Brunner, I., Widmer, F., Zeyer, J., & Frey, B. (2015). Vertical distribution of the soil microbiota along a successional gradient in a glacier forefield. *Molecular Ecology* , 24 (5), 1091–1108. doi: 10.1111/mec.13051
- Rosero, P., Crespo-Perez, V., Espinosa, R., Andino, P., Barragan, A., Moret, P., ... Cauvy-Fraunie, S. (2021). Multi-taxa colonisation along the foreland of a vanishing equatorial glacier. *Ecography* , 44 (7), 1010–1021. doi: 10.1111/ecog.05478
- Sadaka, N., & Ponge, J. F. (2003). Soil animal communities in holm oak forests: influence of horizon, altitude and year. *European Journal of Soil Biology* , 39 (4), 197–207. doi: 10.1016/J.EJSOBI.2003.06.001
- Schaetzl, R., & Anderson, S. (2005). *Soils: Genesis and Geomorphology*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511815560
- Schutte, U. M. E., Abdo, Z., Bent, S. J., Williams, C. J., Schneider, G. M., Solheim, B., & Forney, L. J. (2009). Bacterial succession in a glacier foreland of the High Arctic. *The ISME Journal* 2009 3:11 , 3 (11), 1258–1268. doi: 10.1038/ismej.2009.71
- Schweizer, S. A., Hoeschen, C., Schluter, S., Kogel-Knabner, I., & Mueller, C. W. (2018). Rapid soil formation after glacial retreat shaped by spatial patterns of organic matter accrual in microaggregates. *Global Change Biology* , 24 (4), 1637–1650. doi: 10.1111/GCB.14014
- Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods* , 11 (1), 54–71. doi: 10.1037/1082-989X.11.1.54

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). Environmental DNA: For Biodiversity Research and Monitoring. In *Environmental DNA: For Biodiversity Research and Monitoring*. doi: 10.1093/oso/9780198767220.001.0001

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards Next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21 (8), 2045–2050. doi: 10.1111/j.1365-294X.2012.05470.x

Wickham, H. (2016). *ggplot2*. doi: 10.1007/978-3-319-24277-4

Wietrzyk-Pelka, P., Rola, K., Szymański, W., & Wegrzyn, M. H. (2020). Organic carbon accumulation in the glacier forelands with regard to variability of environmental conditions in different ecogenesis stages of High Arctic ecosystems. *Science of The Total Environment*, 717, 135151. doi: 10.1016/J.SCITOTENV.2019.135151

Zhao, H., Zheng, W., Zhang, S., Gao, W., & Fan, Y. (2021). Soil microbial community variation with time and soil depth in Eurasian Steppe (Inner Mongolia, China). *Annals of Microbiology 2021 71:1*, 71 (1), 1–12. doi: 10.1186/S13213-021-01633-9

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., ... Taberlet, P. (2019, April 1). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, Vol. 28, pp. 1857–1862. Blackwell Publishing Ltd. doi: 10.1111/mec.15060

Zinger, L., Taberlet, P., Schimann, H., Bonin, A., Boyer, F., De Barba, M., ... Chave, J. (2019). Body size determines soil community assembly in a tropical forest. *Molecular Ecology*, 28 (3), 528–543. doi: 10.1111/mec.14919

Zumsteg, A., Luster, J., Göransson, H., Smittenberg, R. H., Brunner, I., Bernasconi, S. M., ... Frey, B. (2011). Bacterial, Archaeal and Fungal Succession in the Forefield of a Receding Glacier. *Microbial Ecology 2011 63:3*, 63 (3), 552–564. doi: 10.1007/S00248-011-9991-8

## DATA ACCESSIBILITY

Raw sequencing data as well as filtered data will be deposited in DRYAD (Dryad Digital Repository) upon acceptance.

## AUTHOR CONTRIBUTIONS

GFF, JP, WT, PT, GAD, DF, RA and MC designed the experiment. SM, GFF, RA, MC, FG, MG, conducted the field work. AG, AB and LG conducted molecular analyses. AG and AB performed the bioinformatic treatment of sequence data. CC, AZ, JP performed edaphic analyses. AG, AC, IC ran statistical analyses. AG, AC and GFF drafted the manuscript. All the authors contributed substantially to the revision process, and accepted the final version.

## ORCID

Alessia Guerrieri: <https://orcid.org/0000-0002-1519-3517>

Alexis Carteron: <https://orcid.org/0000-0003-4162-9254>

Aurélié Bonin: <https://orcid.org/0000-0001-7800-8609>

Ludovic Gielly: <https://orcid.org/0000-0001-5164-6512>

Gentile Francesco Ficetola: <https://orcid.org/0000-0003-3414-5155>

Mauro Gobbi: <https://orcid.org/0000-0002-1704-4857>

Fabrizio Gili: <https://orcid.org/0000-0002-1817-0193>

Pierre Taberlet: <https://orcid.org/0000-0002-3554-5954>

Jérôme Poulenard : <https://orcid.org/0000-0003-0810-0308>

Roberto Ambrosini: <https://orcid.org/0000-0002-7148-1468>

Isabel Cantera: <https://orcid.org/0000-0003-3161-1878>

Silvio Marta: <https://orcid.org/0000-0001-8850-610X>

## TABLE

Table 1. Differences in community structures (Hellinger distances) of the six taxonomic groups across time, depth and their interaction using permutational multivariate analysis of variance (PERMANOVA) and tested for the homogeneity of multivariate dispersions.  $p$  -values were determined using 9999 permutations.

Taxonomic group	Variables	Variance explained (partial $R^2$ )	$p$	Dispersion homogeneity ( $p$ )
Bacteria	Time	0.041	<b>0.001</b>	<b>&lt;0.001</b>
	Depth	0.005	<b>0.001</b>	0.635
	Time $\times$ Depth	0.003	0.102	-
Eukaryota	Time	0.018	<b>0.001</b>	<b>&lt;0.001</b>
	Depth	0.004	<b>0.001</b>	0.173
	Time $\times$ Depth	0.003	0.685	-
Mycota	Time	0.030	<b>0.016</b>	<b>&lt;0.001</b>
	Depth	0.002	<b>0.003</b>	0.394
	Time $\times$ Depth	0.002	0.388	-
Collembola	Time	0.019	0.538	0.104
	Depth	0.003	0.580	0.324
	Time $\times$ Depth	0.004	0.457	-
Insecta	Time	0.021	0.096	<b>&lt;0.001</b>
	Depth	0.003	<b>0.041</b>	0.784
	Time $\times$ Depth	0.003	0.338	-
Oligochaeta	Time	0.027	0.101	<b>0.004</b>
	Depth	0.003	0.128	0.988
	Time $\times$ Depth	0.004	0.157	-

## FIGURES CAPTIONS

Fig. 1. Study area and sampling design. Maps are pseudo-color representation of altitude (source: 30 arc-sec digital elevation model) and land cover (Source: Copernicus Sentinel data 2019) created by AG, SM and GFF.

Fig. 2. Variation of alpha-diversity (measured using  $q = 1$ ) through time and depth in six taxonomic groups. Results obtained using  $q = 0$  can be seen in Table S3.

Fig. 3. Differences in total beta diversity between soil depths through time.

Fig. 4. Ordination of the community structures (Hellinger distances dissimilarities) of the six taxonomic groups in the five proglacial plains at two sampling depths (0-5 cm and 7.5-20 cm). The first two axes of the distance-based Principal Component Analyses are displayed with corresponding percentage of explained variance. Sample points are displayed with color representing time since glacier retreat.







