The development of terrestrial ecosystems emerging after glacier retreat

Gentile Francesco Ficetola, Silvio Marta, Alessia Guerrieri, Isabel Cantera, Aurélie Bonin, Sophie Cauvy-Fraunié, Roberto Ambrosini, Marco Caccianiga, Fabien Anthelme, Roberto Sergio Azzoni, et al.

The development of terrestrial ecosystems emerging after glacier retreat

Gentile Francesco Ficetola[1,2], Silvio Marta[1,3], Alessia Guerrieri[1,4], Isabel Cantera[1], Aurélie Bonin[1,4], Sophie Cauvy-Fraunié[5], Roberto Ambrosini[1], Marco Caccianiga[6], Fabien Anthelme[7], Roberto Sergio Azzoni[1,8], Peter Almond[9], Pablo Alviz Gazitúa[10], Jorge Luis Ceballos Lievano[11], Pritam Chand[12], Milap Chand Sharma[13], John J. Clague[14], Justiniano Alejo Cochachín Rapre[15†], Chiara Compostella[8], Rolando Cruz Encarnación[15], Olivier Dangles[16], Philip Deline[17], Andre Eger[18], Sergey Erokhin[19], Andrea Franzetti[20], Ludovic Gielly[2], Fabrizio Gili[1,21], Mauro Gobbi[22], Sigmund Hågvar[23,24], Rüdiger Kaufmann[25], Norine Khedim[17], Rosa Isela Meneses[26,27], Marco Aurelio Morales-Martínez[28], Gwendolyn Peyre[29], Francesca Pittino[20,30], Angela Proietto[1], Antoine Rabatel[31], Katrin Sieron[28], Levan Tielidze[32,33], Nurai Urseitova[19], Yan Yang[34], Vitalii Zaginaev[19], Andrea Zerboni[8], Anaïs Zimmer[35], Guglielmina Adele Diolaiuti[1], Pierre Taberlet[2,24], Jerome Poulenard[17], Diego Fontaneto[36,37], Wilfried Thuiller[2], Alexis Carteron[38,1]

1. Dipartimento di Scienze e Politiche Ambientali, Università degli Studi di Milano, via Celoria 10, 20133 Milano, Italy
2. Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, F-38000 Grenoble, France
3. CNR - Institute of Geosciences and Earth Resources, Via G. Moruzzi 1, 56124 Pisa, Italy
4. Argaly, Bâtiment CleanSpace, 354 Voie Magellan, 73800 Sainte-Hélène-du-Lac, France
5. INRAE, UR RIVERLY, Centre de Lyon-Villeurbanne, Villeurbanne, France
6. Dipartimento di Bioscienze, Università degli Studi di Milano, via Celoria 26, 20133 Milano, Italy
7. AMAP, Univ. Montpellier, IRD, CIRAD, CNRS, INRAE, Montpellier, France
8. Dipartimento di Scienze della Terra "Ardito Desio", Università degli Studi di Milano, Via L. Mangiagalli 34, 20133 Milano, Italy
9. Department of Soil and Physical Sciences, Lincoln University, Lincoln 7647, New Zealand
10. Departamento de Ciencias Biológicas y Biodiversidad, Universidad de Los Lagos, Osorno, Chile
11. Instituto de Hidrología, Meteorología y Estudios Ambientales IDEAM, Bogotá, Colombia
12. Department of Geography, School of Environment and Earth Sciences, Central University of Punjab, VPO- Ghudda, Bathinda -151401, Punjab, India
13. Centre for the Study of Regional Development - School of Social Sciences, Jawaharlal Nehru University, New Mehrauli Road, 110067 New Delhi, India
14. Department of Earth Sciences, Simon Fraser University, 8888 University Drive, Burnaby, V5A 1S6, Canada
15. Área de Evaluación de Glaciares y Lagunas, Autoridad Nacional del Agua, Huaraz, Peru
16. CEFE, Univ Montpellier, CNRS, EPHE, IRD, Univ Paul Valéry Montpellier 3, Montpellier, France
17. Univ. Savoie Mont Blanc, Univ. Grenoble Alpes, EDYTEM, F-73000 Chambéry, France
18. Mannaki Whenua - Landcare Research, Soils and Landscapes, 54 Gerald St, Lincoln 7608, New Zealand
19. Institute of Water Problems and Hydro-Energy, Kyrgyz National Academy of Sciences, 720033, Frunze, 533, Bishkek, Kyrgyzstan
20. Department of Earth and Environmental Sciences (DISAT) - University of Milano-Bicocca, Milano, Italy
21. Department of Life Sciences and Systems Biology, University of Turin, Via Accademia Albertina 13, 10123 Turin, Italy
22. Research and Museum Collections Office, Climate and Ecology Unit, MUSE-Science Museum, Corso del Lavoro e della Scienza, 3, 38122 Trento, Italy
23. Faculty of Environmental Sciences and Natural Resource Management (INA), Norwegian University of Life Sciences, Universitetstunet 3, 1433 Ås, Norway
24. UiT – The Arctic University of Norway, Tromsø Museum, Tromsø, Norway
25. Department of Ecology, University of Innsbruck, Innsbruck, Austria
26. Herbario Nacional de Bolivia: La Paz, La Paz, Bolivia
27. Universidad Católica del Norte, Antofagasta, Chile
28. Universidad Veracruzana, Centro de Ciencias de la Tierra, CP 91090, Xalapa, Veracruz, Mexico
29. Department of Civil and Environmental Engineering, University of the Andes, Bogotá, Colombia
30. Swiss Federal Institute for Forest, Snow and Landscape Research. Zürcherstrasse 111, 8903 Birmensdorf
31. Univ. Grenoble Alpes, CNRS, IRD, INRAE, Grenoble-INP, Institut des Géosciences de l'Environnement (IGE, UMR 5001), F-38000 Grenoble, France
32. Securing Antarctica's Environmental Future, School of Earth, Atmosphere and Environment, Monash University, Clayton, VIC, 3800, Australia
33. School of Natural Sciences and Medicine, Ilia State University, Cholokashvili Ave 3/5, Tbilisi 0162, Georgia
34. Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu, 610041, China
35. Department of Geography and the Environment, University of Texas at Austin, 78712 Austin, TX, USA
36. CNR - Water Research Institute, Largo V. Tonolli 50, 28922 Verbania, Italy
37. NBFC - National Biodiversity Future Center, Piazza Marina 61, 90133 Palermo, Italy
38. Université de Toulouse, École d'Ingénieurs de PURPAN, UMR INRAE-INPT DYNAFOR, 31076 Toulouse, France

Abstract. The global retreat of glaciers is dramatically altering mountain and high-latitude landscapes, with new ecosystems developing from apparently barren substrates[1–4]. The study of these emerging ecosystems is critical to understanding how climate change interacts with microhabitat and biotic communities and determines the future of ice-free terrains[1,5]. Here, using a comprehensive characterization of ecosystems (soil properties, microclimate, productivity and biodiversity by environmental DNA metabarcoding[6]) across 46 proglacial landscapes worldwide, we found that all the environmental properties change with time since glaciers retreated, and temperature modulates the accumulation of soil nutrients. The richness of bacteria, fungi, plants and animals increases with time since deglaciation, but their temporal patterns differ. Microorganisms colonized most rapidly in the first decades after glacier retreat, while most macroorganisms took longer. Increased habitat suitability, growing complexity of biotic interactions and temporal colonization all contribute to the increase of biodiversity over time. These processes also modify community composition for all the groups of organisms. Importantly, plant communities show positive links with all other biodiversity components and play a key role in ecosystem development. These unifying patterns provide new insights into the early dynamics of deglaciated terrains and highlight the need for integrated surveillance of their multiple environmental properties[5].

Introduction

The global retreat of glaciers is one of the most emblematic signs of the ongoing climate change, exposing increasingly large ice-free areas that are becoming a critical component of mountain and high-latitude landscapes[4,7,8]. These proglacial landscapes rapidly become home to a myriad of organisms, featuring unique biodiversity components and also providing critical ecosystem services[9,10]. Accelerated glacier retreat, coupled with rising temperatures, is leading to a complete restructuring of mountain environments where almost barren substrates, hosting a few resistant or pioneer organisms[11–13], can quickly evolve into complex and biodiverse ecosystems such as alpine meadows and forests. Recognized patterns of ecosystem formation include the development of soils – mostly originating from bedrock or glacial sediments – that support novel ecosystems, an increase in community richness (alpha-diversity), and changes in species composition over time (beta-diversity)[1,2,14].

Identifying the processes that drive and shape the formation of proglacial landscapes has been a central topic in ecology since its inception. These ecosystems are now facing exceptional changes, which reinforces the need to understand the mechanisms that govern their development[5,15,16]. The study of chronosequences, as naturally produced by glacier retreat, provides a unique opportunity to reconstruct ecosystem formation by comparing areas in a similar context but with different times since deglaciation[17,18]. Previous local studies have proposed three key processes driving the formation of biological communities after glacier retreat: habitat filtering, biotic interactions, and processes mediated by time[1,19,20]. The habitat filtering hypothesis predicts that the build-up of communities is strongly dependent on the availability of suitable microhabitats, such as nutrient-rich soils or favorable microclimates. The biotic interaction hypothesis emphasizes the importance of relationships among organisms and predicts that certain organisms (e.g. plants, soil microorganisms) play a key role in enabling the establishment of organisms at higher trophic levels. Finally, the time hypothesis suggests that the probability of dispersal and colonization by propagules increases with time since glacier retreat. The latter predicts an increase in community richness over time that cannot be attributed to changes in habitat and biotic interactions[1,21]. However, there is still no consensus on the actual role of these three processes[14,22], perhaps because their importance may vary between lifeforms and regions of the world. The lack of standardized data, measuring multiple environmental properties and biodiversity components across diverse geographical areas, has so far hampered the development of a comprehensive framework needed to explicitly assess the processes underlying the dynamics of biotic and abiotic components of these emerging ecosystems.

2

In this study, we took on the challenge of building a broad-scale dataset to assess the processes underlying the formation of terrestrial ecosystems following the retreat of mountain and high-latitude glaciers. We did this by conducting a comprehensive inventory of 1,251 plots along 46 chronosequences in proglacial landscapes from tropical, temperate and subpolar regions, with sites ranging from 1 to 483 years since glacier retreat (Extended Data Fig. 1). Ecosystem formation was measured using three main properties: i) soil properties and nutrients (pH, nitrogen, carbon and phosphorus); ii) vegetation primary production, and iii) local-scale biodiversity. Biodiversity was measured using environmental DNA (eDNA) metabarcoding to provide information on eight groups ranging from microorganisms (bacteria, fungi and protists) to macroorganisms (vascular plants and animals, with a particular focus on earthworms, insects and springtails; Fig. 1, Supplementary Tables S1-S2)[6]. These properties strongly influence ecosystem functions[17,23], and a better knowledge of their temporal dynamics and of the processes involved is a prerequisite for the management of these emerging environments[3,5]. Specifically, we ask: What are the temporal dynamics of environmental properties since glacier retreat? How are these environmental features related to each other? What is the importance of temporal colonization, habitat filtering and biotic interactions in the formation of biotic communities? Are these processes valid across climatic regions, or are there differences between them?

Dynamics of environmental properties

Our broad-scale assessment shows that all the environmental properties vary strongly with time since glacier retreat but follow distinct temporal trajectories. Concentrations of soil nutrients and organic matter (phosphorus, nitrogen and carbon) and their stoichiometric ratios (C/N and C/P) significantly increased during early ecosystem establishment, with a particularly rapid growth in later successional stages (Fig. 1A, Extended Data Table 1). Although soils developing after glacier retreat are generally thought to be nutrient-poor[2,24], in some landscapes we detected a significant concentration of nutrients immediately after glacier retreat (e.g. Agola and Forni in the Central Alps). These nutrients may have been derived from mountain slopes, supraglacial debris, ancient organic matter buried in or under the ice, or from glacier and snow microbes[24–27], and are likely essential for triggering the formation of early-stage communities of microbes and other pioneer organisms[28] (Fig. 1). Although variable across proglacial landscapes, pH decreased slowly over time (Fig. 1a), except in the first years after glacier retreat[2]. Progressing from bare-ground to later successional stages, ecosystem production, estimated by the Normalized Difference Vegetation Index (NDVI), increased from near-zero at sites less than 10 years old (NDVI < 0.2) to higher values at the oldest sites.

Metabarcoding of soil eDNA yielded 10,825 molecular operational taxonomic units (MOTUs; see Methods section) globally, allowing us to characterize the biodiversity of the major taxonomic groups living in, or associated with, the soil. Across all sites, the total number of MOTUs detected was highly variable among taxonomic groups, ranging from 331 (earthworms) to > 3000 (bacteria and fungi; Supplementary Table S1). The local diversity (alpha-diversity) of all groups of organisms was positively correlated and strongly increased with time since glacier retreat, still there were clear differences among taxa (Fig. 1b; Extended Data Fig. 2). Immediately after glacier retreat (< 10 years, Fig. 1b), the communities of bacteria and, to a lesser extent, fungi and protists were already diverse, whereas plant and animal diversity was more limited, as these organisms were extremely scarce in the youngest plots. Early communities can include taxa originating from the glacier surface or under the ice, but also efficient dispersers that are early colonizers[11–13]. During the early stages of ecosystem formation, the local diversity of all groups increased sharply. Bacterial and fungal diversity increased faster during the first ~50-100 years after glacier retreat than in the later phases, as indicated by the negative non-linear trend of the temporal changes (Fig. 1b, right panels). In contrast, for vascular plants, earthworms, and springtails, local diversity increased slowly until 10-20 years after glacier retreat, followed by a sharp increase (positive non-linear term; Fig. 1b; Supplementary Table S1), with no sign of plateauing or deceleration over the entire period of our

records. The predictions of these models matched the development of diversity observed in independent studies using permanent plots (Extended Data Fig. 3), suggesting that our chronosequence-based analyses provide a good representation of biodiversity changes occurring through time.
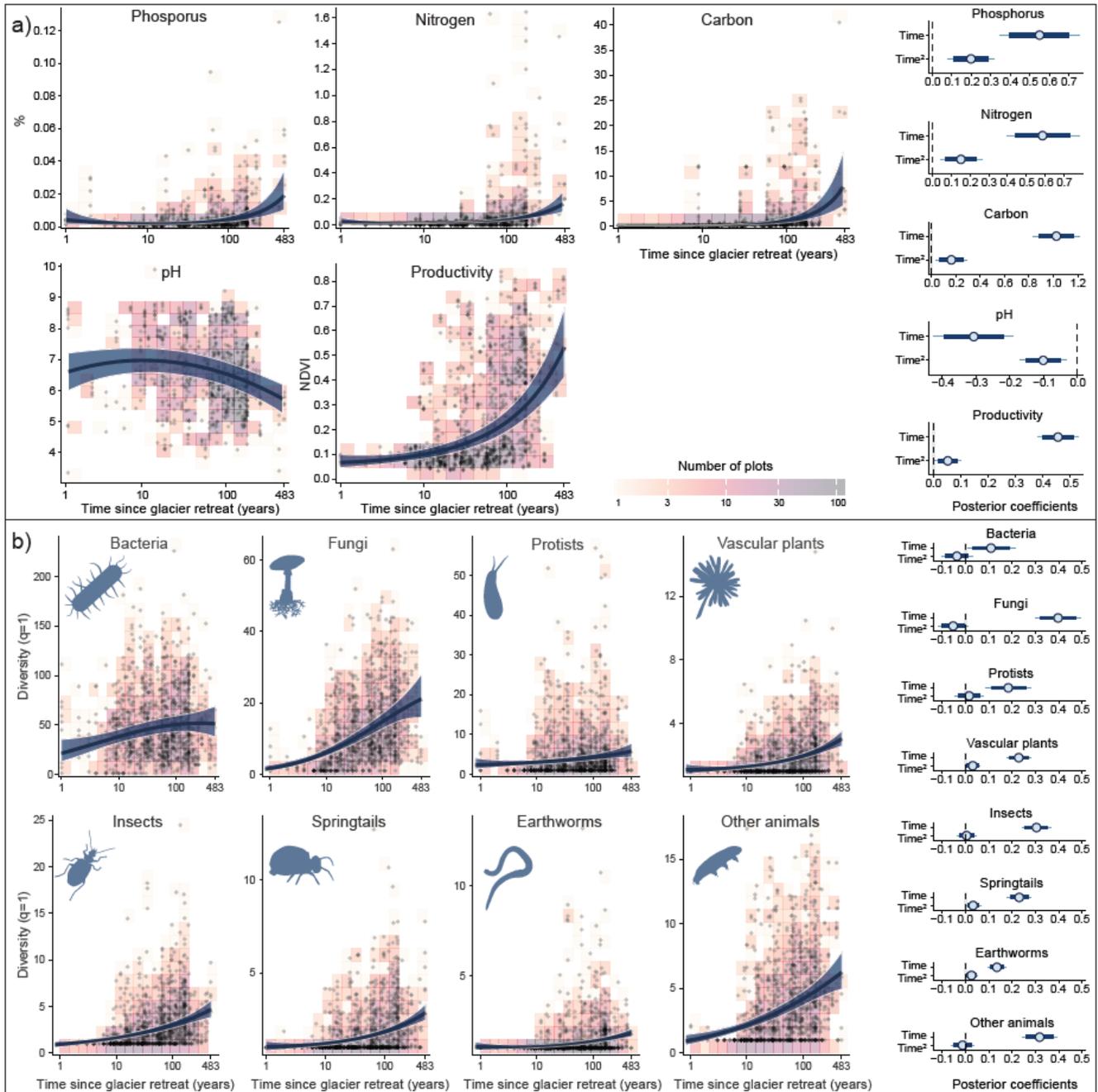
Figure 1. Development of environmental properties with time since glacier retreat. a) Soil nutrients (phosphorus, nitrogen), organic carbon, pH and productivity. Productivity is estimated from the normalized difference vegetation index (NDVI; values < 0.2 indicate very low productivity or even barren areas). For pH, the quadratic term overlapped zero after removing the plots with time since glacier retreat < 10 years (Extended Data Table 1). b) Alpha-diversity (Hill's number $q = 1$) of the eight groups. In the left panels, bold lines indicate the regression curves obtained through Bayesian generalized linear mixed models; shaded areas represent the 95% credible intervals of the regression. The right panels show the posterior coefficient estimates for the effects of time (linear and quadratic terms). Circles represent medians; bold and thin lines represent 95% and 99% credible intervals of coefficients, respectively. See Extended Data Table 1 for detailed results and sample size.

In some taxa (bacteria, plants, insects, springtails, and earthworms), colonization dynamics differed across areas of the world, being faster in temperate areas compared to subpolar and tropical regions (Extended Data Fig. 4). This probably occurred because warm temperatures during the growing season in temperate regions (Extended Data Fig. 4) accelerate the development of several environmental properties that favor the establishment of these organisms, such as productivity and soil nutrients[24] (see below). In addition, most of the tropical sites were at extreme altitudes, where geomorphic processes and disturbance limit the accumulation and persistence of soil organic matter[29,30].

Drivers of ecosystem development

We used a structural equation model (SEM) approach to identify relationships among environmental properties and to test the importance of the passage of time since glacier retreat, microhabitat (soil pH, nutrients, productivity and microclimate, i.e., soil temperature and wetness) and biotic interactions on the temporal changes in local diversity. The relationships between local diversity and some environmental properties, such as productivity and soil nutrients, can be complex and may involve feedbacks[1,31], so we first tested the support of three alternative models: *1*) a model in which soil nutrients and ecosystem productivity determine the diversity of soil organisms; *2*) an inverse model in which the diversity of all the organisms determines soil nutrients and productivity; and *3*) a model in which diversity covaries with soil nutrients and productivity during ecosystem formation (see Methods section and Extended Data Figs. 5-6). The covariation model received the best support from the data (difference in Bayesian Information Criterion [BIC] >30; Extended Data Fig. 6, Extended Data Table 2), suggesting that soil abiotic and biotic compartments co-evolve during ecosystem formation, with low diversity and limited productivity in recently deglaciated, cold and nutrient-poor terrains. The interplay and feedback between ecosystem properties and diversity then determine the development of these ecosystems through time (Fig. 2).
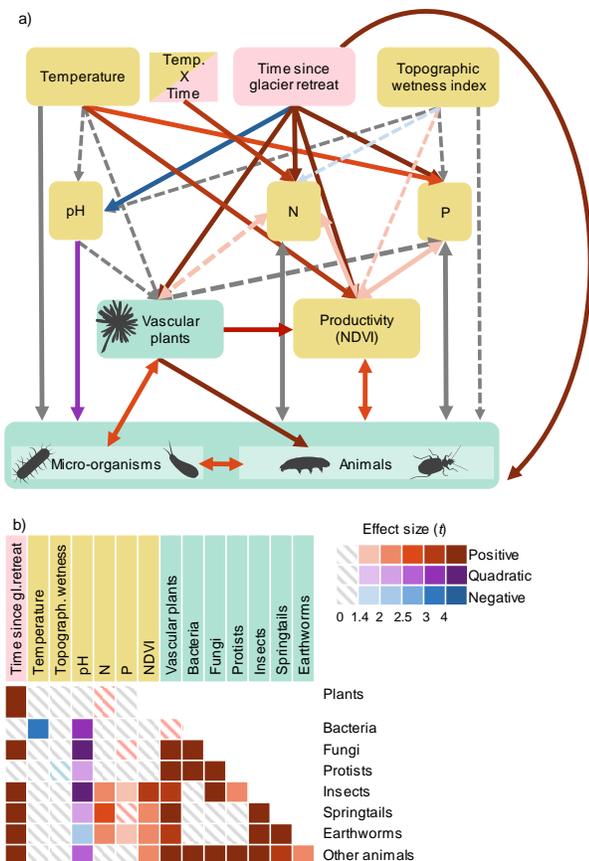
Figure 2. Direct and indirect relationships between properties of ecosystems emerging after the retreat of glaciers. a) Pathways show relationships among time since glacier retreat, microclimate, soil nutrients and pH, ecosystem productivity and alpha-diversity of organisms. Single-head arrows indicate tested causal relationships; double-head arrows indicate covariation. The color of arrows and squares represents the effect size of relationships. For simplicity, in a) we only show one arrow between each environmental property and all the groups analyzed for soil biodiversity; the color of the arrow represents the median of the effect size across these groups; dashed lines indicate relationships that are not significant for any of the considered groups. b) Detailed relationships between each taxonomic group and other variables/other taxa; dashed cells indicate non-significant relationships (see also Supplementary Table S3). Only the best-BIC model is shown here. See Extended Data Fig. 6, Extended Data Table 2 and Supplementary Table S4 for models with alternative structure or that consider carbon instead of nitrogen. $N = 793$ plots.

We therefore used the covariation SEM to quantify the importance of processes likely involved in community development. More explicitly, we tested whether time, microhabitat and biotic interactions had direct or indirect effects on biodiversity. Strong relationships between diversity and microhabitat would support the habitat filtering hypothesis, while strong relationships between the diversity of the different taxonomic groups after accounting for habitat differences would support the biotic interaction hypothesis. Finally, a strong effect of time since glacier retreat, unexplained by habitat or community variation, would support a key role played by dispersal limitation[20]. We showed that soil nutrients increased significantly over time. A warm microclimate had a positive effect on soil nutrients, with a particularly rapid accumulation of nitrogen under warmer conditions[24], as shown by the interaction between time and temperature (Fig. 2; Supplementary Table S3). This probably occurs because warm temperatures increase the activity of N-fixing bacteria[32,33]. Time emerged as the strongest driver of plant diversity, suggesting that dispersal constraints control the increase in plant richness. After accounting for the effect of time, plant diversity also showed a positive covariation with fungal diversity and, to a lesser extent, soil nitrogen. Local microbial diversity was mainly related to plant diversity and soil pH. Both acidic and alkaline conditions limited microbial diversity, as indicated by quadratic trends in the models. pH has direct effects on the availability of several cations and is a predominant driver of microbial diversity at both local and continental scales[34]. Furthermore, the diversity of most microorganisms covaried over time, supporting a pervasive role for biotic interactions. The diversity of most animals responded positively to plant diversity and to microhabitat features such as soil nutrients and ecosystem productivity, which indicate resource-rich ecosystems. Covariation in diversity was also strong for all the animal groups (Fig. 2, Supplementary Tables S3-S4). However, even after accounting for these interdependencies, time remained a major predictor of the diversity for all organisms except bacteria and protists.
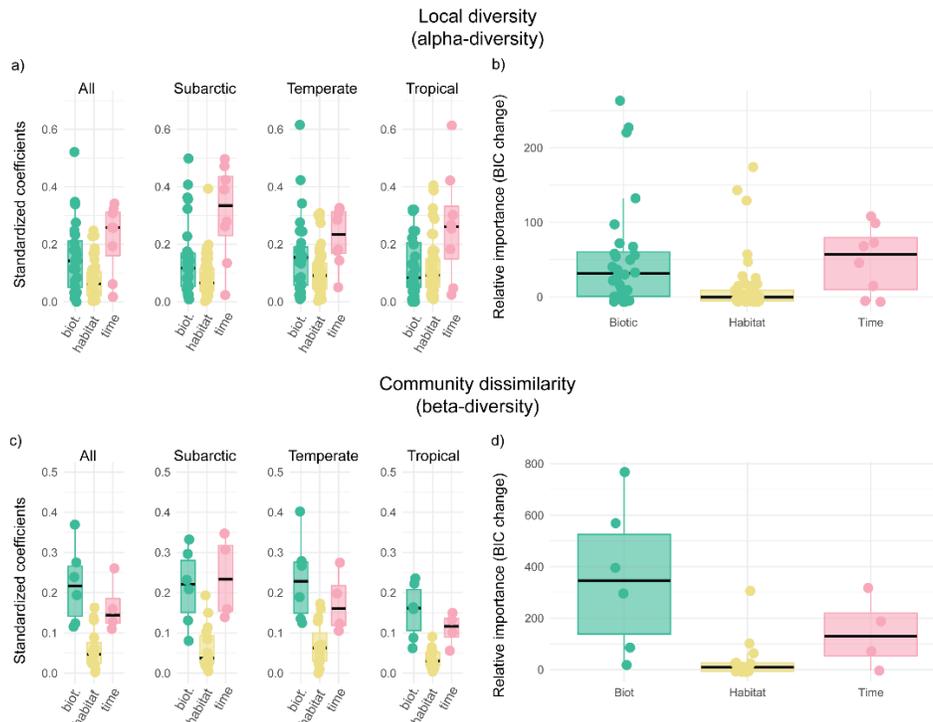
Figure 3. Importance of biotic relationships, habitat and time for biodiversity development. In a) and c), the importance of candidate drivers was measured as the standardized coefficients of paths representing biotic relationships, habitat, and time since glacier retreat. Coefficients were obtained from structural equation models (SEMs) assessing the drivers of: a) local diversity (alpha-diversity) and c) community dissimilarity (beta-diversity). First, we ran a SEM using all the proglacial landscapes, and then only considering subpolar, temperate and tropical forelands. In b) and d), importance was measured as the loss of performance (BIC value) when paths representing biotic relationships, habitat and time since glacier retreat are removed from SEMs. In a) and b), we show the effects of 89 alternative paths; in c) and d) we show the effects of 25 paths. Bold lines indicate the median value, box limits comprise the 25th and 75th percentile while whiskers extend up to 1.5 × interquartile range from the box limits.

The importance of the three processes we considered (habitat filtering, biotic interactions and time-mediated colonization) was fairly consistent across climatic regions, and in all of them time was proportionally more important than biotic interactions and habitat. The effect of time was particularly strong in subpolar landscapes, compared to those of temperate and tropical regions (Fig. 3a). Subpolar forelands are covered with snow for most of the year, and this can exacerbate the difficulties of dispersal and colonization. Nevertheless, all three processes played a significant role in the formation of multi-trophic communities and ecosystems, as deleting SEM paths representing any of these processes resulted in unsupported models (Fig. 3b).
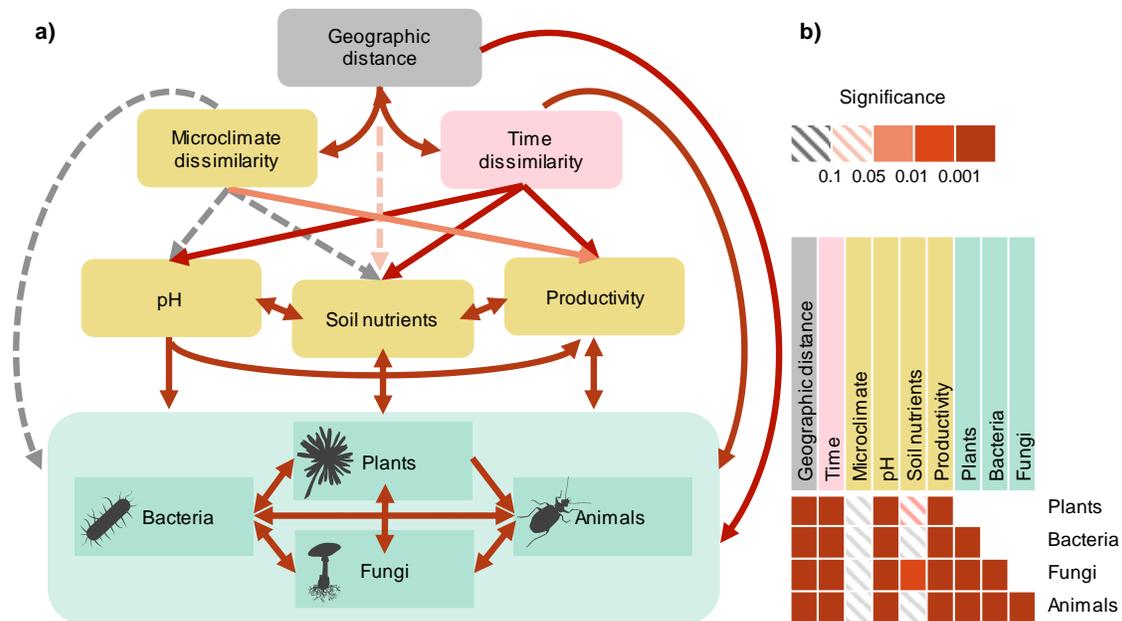

Figure 4. Direct and indirect effects of ecosystem properties on the dissimilarity of communities. a) Pathways show relationships among time difference, geographic distance, dissimilarity of ecosystem properties (microclimate, productivity, soil nutrients and pH) and the beta-diversity of soil organisms. Single-head arrows indicate tested causal relationships; double-head arrows indicate covariation. For simplicity, in a) we only show one arrow for all the taxonomic groups. Animals combine the animal taxa detected by all the markers. b) Detailed relationships between each taxonomic group and other variables/other taxa (see also Supplementary Table S5). The colors of arrows represent the significance of relationships, assessed using 9,999 permutations. $N = 5741$ pairs of plots within glacier forelands.

Drivers of community composition

Local diversity is only a partial descriptor of overall biodiversity, as local communities with similar diversity or richness may have very different compositions. The extent to which community composition changes over time within a glacier foreland is questionable. Such community dissimilarities may have important implications for the long-term coexistence of meta-communities (e.g. spatial storage effects[35]) and for the protection of biodiversity across entire landscapes[5]. We therefore tested whether the community dissimilarity between plots of the same proglacial landscape was driven by the same processes that affect local diversity. Again, we built a SEM to disentangle the direct and indirect effects of environmental features and time since glacier retreat as drivers of community dissimilarity. To provide integrative measures of environmental variation, we also calculated the total dissimilarity of microclimate (based on temperature and topographic wetness) and of soil nutrients (carbon, nitrogen and phosphorus) between plots, assuming covariation between the dissimilarity of soil communities, soil nutrients and productivity.

Temporal differences between plots within the same proglacial landscape had strong impacts on the dissimilarity of soil properties and productivity, and on the community dissimilarity of all soil organisms (Fig. 4). As expected, differences in soil pH strongly influenced the community composition of all the organisms, whereas

microclimatic differences within the same landscape had little effect, probably because the limited size of forelands reduced the variation of climatic conditions. As with local diversity, the covariation between the community dissimilarity of different organisms was also very high. Nutrients co-varied with the communities of fungi and plants, although this latter relationship was not strong (Fig. 4, Supplementary Table S5).

Biotic interactions were evident between all the groups of organisms and emerged as the primary drivers of the evolution of community composition. Nevertheless, our results support a joint role of habitat formation, biotic interactions and time[36], as removing relationships representing any of these three processes dramatically reduced model performance (Fig. 3c-d). Also in this case, coefficients obtained for distinct geographical regions were similar to the ones obtained using all the data, with time showing slightly higher importance in subpolar landscapes compared to the temperate and tropical ones (Fig. 3c).

Complex evolution of deglaciated areas

Our broad-scale analysis showed that many features of the ecosystems that emerge after glaciers retreat evolve together. Time since glacier retreat, habitat development and biotic interactions jointly affect the establishment of organisms, but each process has a distinct role on the different facets of biodiversity. Time since glacier retreat can have both direct and indirect effects on community richness and composition. Indirect effects include the weathering of exposed substrates and the accumulation of autochthonous and allochthonous organic matter over time, that trigger soil development and favor the formation of microhabitats for many organisms, such as detritivorous animals (e.g. earthworms and springtails; Fig. 2). Nevertheless, for most organisms, time remained the strongest predictor of diversity even after accounting for habitat variation. This suggests that community formation is heavily affected by dispersal-related processes, such as the increased probability of colonization by propagules or moving individuals [21,37]. Time is less important for the richness of bacteria and protists, which are less dispersal-limited than larger organisms[38] and are present immediately after glacier retreat[34].

Covariation between different organisms is central to understanding the evolution of both community diversity and composition within proglacial landscapes. Relationships are particularly strong among microorganisms (bacteria, fungi and protists), among the different animal taxa, and between plants and both microorganisms and animals. In principle, strong relationships between biotic components do not necessarily imply direct interactions, as they could be caused by covariation with unmeasured variables[39], still we integrated in our analyses some of the most widely recognized drivers of soil biodiversity to limit the impact of unmeasured environmental variation. Our models suggest a key role for positive interactions[40], such as facilitation and increased food availability for organisms at higher trophic levels, for instance ground beetles preying on springtails and nematodes feeding on bacteria or fungi[41,42]. Vascular plants are at the center of the interaction network, as they were the only taxon to show clear relationships with all other organisms. First, plants interact strongly with many soil microorganisms (bacteria and fungi), which develop rather rich communities immediately after glacier retreat[43–46]. Then, once established, plants become crucial for animal colonization, providing food, shelter and even substrate for many animals and determining ecosystem productivity[14].

Over time, the diversity and composition of taxa undergo substantial changes, with each group responding distinctively to various factors aligned with their respective life histories. Fungal and bacterial communities show early diversity post-glacier retreat, then plateauing around a century later. This may be attributed to the peak of diversity of some microorganism groups 50-100 years after the onset of soil development, followed by the replacement of pioneer, oligotrophic taxa with copiotrophic ones[1,2,20]. Alternatively, the vegetation succession from open habitats to closed forests can lead to the replacement of microorganisms associated to early colonizer plants by the ones associated to the late-successional ones[20,44,47]. For larger soil organisms, dispersal limitation and habitat play a strong filtering role, requiring more time for initial colonization. Non-linear trends emerge, with significant increases in colonization rates for various animals over time (Fig. 1), likely influenced by the direct

8

and indirect effects of time, habitat changes, and biotic interactions. For instance, time increases the probability of colonization by springtails and earthworms (direct effect). Time also increases the organic matter and nutrients available in soil (habitat-mediated effect) and the richness and productivity of plant communities, which are an important food source for these animals, with potential positive feedbacks[48]. The long time required for ecosystem development thus results from the complex interplay among the multiple components of the community, with different colonization patterns among taxa, and many organisms requiring the presence of other organisms that promote habitat modification or with which they interact. Despite variations in the effect of specific drivers across taxa, the model assuming a joint role of time, habitat and biotic interactions held for all the geographic regions, suggesting a widespread effect of these ecological processes.

Study limitations include the incomplete sampling in some world regions (particularly in subpolar and polar regions), the reliance on chronosequences, and the use of metabarcoding data that lack biomass information and are not fully comparable with traditional inventories. However, the striking congruence between metabarcoding and traditional data (Extended data Fig. 2), coupled with consistent support for fundamental processes across diverse regions worldwide, fortifies the robustness of our findings. Future endeavors aimed at harmonizing metabarcoding and field-collected functional data would facilitate a more direct evaluation of ecosystem functioning and services[49]. Additionally, it is imperative to investigate the role of vertebrates, which can exert keystone impacts such as seed dispersal, soil disturbance, and acceleration of nutrient accumulation and cycling[50,51].

Proglacial ecosystems undergo unparalleled transformations over time, marked by soil development and significant alterations in terrestrial biodiversity resulting from the interplay between habitat characteristics, organism interactions, and dispersal-related dynamics. Given the complex relationships among various components of these ecosystems, efficient monitoring of their transformations requires interdisciplinary approaches that integrate ecological, genetic, and environmental data. Through comprehensive field studies complemented by advanced modeling techniques, we can anticipate the structural changes in these emerging ecosystems, as well as the swift evolution of the services they provide[3,5].

MAIN REFERENCES

1. Ficetola, G. F. *et al.* Dynamics of ecological communities following current retreat of glaciers. *Annu. Rev. Ecol. Evol. Syst.* **52**, 405–426 (2021).
2. Pothula, S. K. & Adams, B. J. Community assembly in the wake of glacial retreat: A meta-analysis. *Global Change Biology* **28**, 6973–6991 (2022).
3. Bosson, J. B. *et al.* Future emergence of new ecosystems caused by glacial retreat. *Nature* **620**, 562–569 (2023).
4. Rounce, D. R. *et al.* Global glacier change in the 21st century: Every increase in temperature matters. *Science* **379**, 78–83 (2023).
5. Zimmer, A., Beach, T., Klein, J. A. & Recharte Bullard, J. The need for stewardship of lands exposed by deglaciation from climate change. *Wiley Interdiscip. Rev. Clim. Change* **13**, e753 (2022).
6. Taberlet, P., Bonin, A., Zinger, L. & Coissac, E. *Environmental DNA: For Biodiversity Research and Monitoring*. (Oxford University Press, Oxford, 2018). doi:10.1093/oso/9780198767220.001.0001.
7. Hock, R. *et al.* GlacierMIP – A model intercomparison of global-scale glacier mass-balance models and projections. *J. Glaciol.* **65**, 453–467 (2019).
8. Lee, J. R. *et al.* Climate change drives expansion of Antarctic ice-free habitat. *Nature* **547**, 49–54 (2017).
9. Körner, C. Mountain Biodiversity, Its Causes and Function. *Ambio* **33**, 11–17 (2004).
10. Palomo, I. Climate Change Impacts on Ecosystem Services in High Mountain Areas: A Literature Review. *Mt. Res. Dev.* **37**, 179–187 (2017).
11. La Farge, C., Williams, K. H. & England, J. H. Regeneration of Little Ice Age bryophytes emerging from a polar glacier with implications of totipotency in extreme environments. *Proceedings of the National Academy of Sciences* **110**, 9839–9844 (2013).
12. Donhauser, J. & Frey, B. Alpine soil microbial ecology in a changing world. *FEMS Microbiol Ecol* **94**, (2018).
13. Hågvar, S. *et al.* Ecosystem Birth near Melting Glaciers: A Review on the Pioneer Role of Ground-Dwelling Arthropods. *Insects* **11**, 644 (2020).
14. Cauvy-Fraunié, S. & Dangles, O. A global synthesis of biodiversity responses to glacier retreat. *Nat. Ecol. Evol.* **3**, 1675–1685 (2019).
15. Hugonnet, R. *et al.* Accelerated global glacier mass loss in the early twenty-first century. *Nature* **592**, 726–731 (2021).
16. Moore, J. W. *et al.* Mining stakes claim on salmon futures as glaciers retreat. *Science* **382**, 887–889 (2023).
17. Poorter, L. *et al.* Multidimensional tropical forest recovery. *Science* **374**, 1370–1376 (2021).
18. Walker, L. R., Wardle, D. A., Bardgett, R. D. & Clarkson, B. D. The use of chronosequences in studies of ecological succession and soil development. *J. Ecol.* **98**, 725–736 (2010).
19. Connell, J. H. & Slatyer, R. O. Mechanisms of Succession in Natural Communities and Their Role in Community Stability and Organization. *Am. Nat.* **111**, 1119–1144 (1977).
20. Hanusch, M., He, X., Ruiz-Hernández, V. & Junker, R. R. Succession comprises a sequence of threshold-induced community assembly processes towards multidiversity. *Commun. Biol.* **5**, 1–9 (2022).
21. Pulsford, S. A., Lindenmayer, D. B. & Driscoll, D. A. A succession of theories: purging redundancy from disturbance theory. *Biol. Rev.* **91**, 148–167 (2016).
22. Rosero, P. *et al.* Multi-taxa colonisation along the foreland of a vanishing equatorial glacier. *Ecography* **44**, 1010–1021 (2021).
23. Fan, K. *et al.* Soil biodiversity supports the delivery of multiple ecosystem functions in urban greenspaces. *Nat. Ecol. Evol* **7**, 113–126 (2023).
24. Khedim, N. *et al.* Topsoil organic matter build-up in glacier forelands around the world. *Glob. Change Biol.* **27**, 1662–1677 (2021).
25. Lutz, S. *et al.* The biogeography of red snow microbiomes and their role in melting arctic glaciers. *Nat. Commun.* **7**, 11968 (2016).
26. Rime, T., Hartmann, M. & Frey, B. Potential sources of microbial colonizers in an initial soil ecosystem after retreat of an alpine glacier. *ISME J.* **10**, 1625–1641 (2016).
27. Zimmer, A. *et al.* Soil temperature and local initial conditions drive carbon and nitrogen build-up in young proglacial soils in the Tropical Andes and European Alps. *CATENA* **235**, 107645 (2024).
28. Bardgett, R. D. *et al.* Heterotrophic microbial communities use ancient carbon following glacial retreat. *Biol. Lett.* **3**, 487–490 (2007).
29. Hunter, B. D., Roering, J. J., Silva, L. C. R. & Moreland, K. C. Geomorphic controls on the abundance and persistence of soil organic carbon pools in erosional landscapes. *Nat. Geosci.* **17**, 151–157 (2024).
30. Draebing, D., Mayer, T., Jacobs, B. & McColl, S. T. Alpine rockwall erosion patterns follow elevation-dependent climate trajectories. *Commun Earth Environ* **3**, 1–12 (2022).
31. Erhart, H. *La Génèse Des Sols En Tant Que Phénomène Géologique : Esquisse d'une Théorie Géologique et Géochimique : Biostasie et Rhexistasie*. (Masson, Paris, 1951).
32. Salazar, A., Warshan, D., Vasquez-Mejia, C. & Andrésson, Ó. S. Environmental change alters nitrogen fixation rates and microbial parameters in a subarctic biological soil crust. *Oikos* **2022**, e09239 (2022).
33. Sepp, S.-K. *et al.* Global diversity and distribution of nitrogen-fixing bacteria in the soil. *Frontiers in Plant Science* **14**, (2023).
34. Bahram, M. *et al.* Structure and function of the global topsoil microbiome. *Nature* **560**, 233–237 (2018).

35. Angert, A. L., Huxman, T. E., Chesson, P. & Venable, D. L. Functional tradeoffs determine species coexistence via the storage effect. *Proceedings of the National Academy of Sciences* **106**, 11641–11645 (2009).

36. Peyre, G. *et al.* The fate of páramo plant assemblages in the sky islands of the northern Andes. *Journal of Vegetation Science* **31**, 967–980 (2020).

37. Vellend, M. *et al.* Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos* **123**, 1420–1430 (2014).

38. Martiny, J. B. H. *et al.* Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* **4**, 102–112 (2006).

39. Ohlmann, M. *et al.* Mapping the imprint of biotic interactions on β-diversity. *Ecol. Lett.* **21**, 1660–1669 (2018).

40. Tscherko, D., Hammesfahr, U., Zeltner, G., Kandeler, E. & Böcker, R. Plant succession and rhizosphere microbial communities in a recently deglaciated alpine terrain. Basic Appl Ecol 6: 367-383. *Basic and Applied Ecology* **6**, 367–383 (2005).

41. Losapio, G. *et al.* Network motifs involving both competition and facilitation predict biodiversity in alpine plant communities. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2005759118 (2021).

42. Sint, D., Kaufmann, R., Mayer, R. & Traugott, M. Resolving the predator first paradox: Arthropod predator food webs in pioneer sites of glacier forelands. *Mol Ecol* **28**, 336–347 (2019).

43. Bennett, J. A. *et al.* Plant-soil feedbacks and mycorrhizal type influence temperate forest population dynamics. *Science* **355**, 181–184 (2017).

44. Calderón-Sanou, I. *et al.* Cascading effects of moth outbreaks on subarctic soil food webs. *Sci Rep* **11**, 15054 (2021).

45. Houlton, B. Z., Wang, Y.-P., Vitousek, P. M. & Field, C. B. A unifying framework for dinitrogen fixation in the terrestrial biosphere. *Nature* **454**, 327–330 (2008).

46. Tedersoo, L., Bahram, M. & Zobel, M. How mycorrhizal associations drive plant population and community biology. *Science* **367**, (2020).

47. Cantera, I. *et al.* The importance of species addition 'versus' replacement varies over succession in plant communities after glacier retreat. *Nat. Plants* **10**, 256–267 (2024).

48. Pugnaire, F. I. *et al.* Climate change effects on plant-soil feedbacks and consequences for biodiversity and functioning of terrestrial ecosystems. *Science Advances* **5**, eaaz1834 (2019).

49. Guerra, C. A. *et al.* Global hotspots for soil nature conservation. *Nature* 1–6 (2022) doi:10.1038/s41586-022-05292-x.

50. Sytsma, M. L. T., Lewis, T., Bakker, J. D. & Prugh, L. R. Successional patterns of terrestrial wildlife following deglaciation. *J. Anim. Ecol.* **92**, 723–737 (2023).

51. Butler, D. R., Anzah, F., Goff, P. D. & Villa, J. Zoogeomorphology and resilience theory. *Geomorphology* **305**, 154–162 (2018).

METHODS

Sampling

Between 2014 and 2020, we collected a soil composite sample in each of 1251 distinct plots across 265 sites, located in the forefield of 46 mountain and high-latitudes glaciers around the world (hereafter: proglacial landscapes; Extended Data Fig. 1), including regions with different rates of glacier retreat[52]. We focused on proglacial landscapes for which high-quality information on the dates of deglaciation after the end of the Little Ice Age was available based on historical archival records including topographical maps, field measurements, repeated photographs, paintings, remote imaging and field data. A complete description of the approaches used for each foreland is available in ref.[53].

For each landscape, we selected 3-17 sites (mean = 5.8, SD = 2.5), corresponding to the position of the glacier front at a given time (Extended Data Fig. 1), and thus creating a chronosequence. Terrain ages of sampled sites range from 1 to 483 years. The number of sites and their age depended on accessibility and the availability of information on past glacier positions. The selected sites were validated by researchers experienced in the geomorphology and mapping of deglaciated areas of each study region to exclude sites affected by disturbances (erosion, streams, surface instability, grazing and heavy recreational activities), with superimposed deposits of different ages (for example, outwash or landslides), and for which dating was uncertain. Some glaciers can undergo local advances, but we only considered sites that were not affected by advances on the basis of geomorphological evidence and archival data[53].

At each site, we established ~5 plots (mean = 4.7 SD = 0.8) as separate spatial replicates of the same age within sites. Plots were regularly spaced at a distance of 20 m where possible, to ensure robust biodiversity sampling[54]. Sampling occurred during the warmest months (i.e. mid-summer); all plots of the same foreland were sampled during the same week, to avoid seasonal variation. At each plot, we collected five soil subsamples within a 1 m radius (Extended Data Fig. 1). Soil subsamples were collected over a depth range of 0-20 cm, avoiding litter, visible roots and plant material. Immediately after glacier retreat, sediments at depth >10 cm host few lifeforms, but at later stages soils become deeper and eDNA-based analyses revealed little difference on community composition along the first 20 cm of soil[24,55,56]. The five subsamples within each plot were then pooled, resulting in a ~200 g composite sample per plot (total: 1251 composite samples; one distinct composite sample per plot). To avoid contamination, soil collection was performed wearing disposable gloves and facemasks; we used a portable blow torch (>1000°C) to decontaminate tools before the collection of each sample. After homogenization of the composite sample, 15 g of soil were taken and immediately placed in a sterile box with 40 g of silica gel as a desiccant. This method ensures reliable preservation of eDNA; for more details, see ref[57].

Soil features and nutrients

For a subset of 32 proglacial landscapes out of the 46 in total, one additional sample per plot was taken and four chemical properties of soil were measured, representing the microhabitat and nutrients available for soil-related organisms. Landscapes with soil properties were representative of all the sampled continents and of all the climatic zones (Supplementary Table S6), and their environmental features and time since glacier retreat were similar to those of forelands without these data (Supplementary Table S7). We considered soil pH, organic carbon and nutrients (nitrogen and phosphorus). Soil pH was measured using a pH-meter from a suspension composed of 4 g of soil and 10 ml of bi-distilled water. Total organic carbon and nitrogen were measured by elemental analysis using a Flash2000 OEA analyzer (ThermoFisher). Depending on pH, two different methods were used to measure assimilable phosphorus through inductively coupled plasma mass spectrometry (iCAP RQ ICP-MS, ThermoFisher): the Bray and Kurtz method[58] for samples with pH < 6.5 and the Olsen method[59] for samples with pH ≥ 6.5.

Soil microclimate and productivity

As measures of soil microclimate, we calculated the mean soil temperature through the growing season and the topographic wetness index (TWI). High-resolution temperature was reconstructed using a global microclimatic model based on macroclimate, local topography (slope, aspect), solar radiation, distance from glacier front and permafrost occurrence[60]. Temperature data were calibrated on 175 plots from subpolar, tropical and temperate landscapes analyzed in this study, and globally validated using near-subsurface records[60]. This model predicted very well ($R^2 = 0.85$) actual values of soil temperature in our plots[60]. Mean soil temperature during the growing season was calculated as the average temperature of the snow-free months over the 2015-2019 period[60]. Soil moisture was approximated using the topographic wetness index with the *upslope.area* function from the *dynatopmodel* R package[61] based on the ASTER Global Digital Elevation Map v3 (https://doi.org/10.5067/ASTER/ASTGTM.003) with 1 arc-second resolution (~30 m at the equator). We used the normalized difference vegetation index (NDVI) as an indicator of gross primary production. NDVI is known to be positively related to annual above-ground net primary production[62] and is considered a reliable approach to estimate vegetation productivity in mountain regions[63]. Cloudiness and water vapor represent serious issues for indexes based on remotely sensed images due to the technical difficulties to distinguish between snow / ice and clouds; this is particularly relevant for global analyses targeting high-mountain areas. To overcome this limitation while retaining the information on the pattern of primary production within each landscape, we used the average maximum NDVI[64,65]. Yearly maxima were calculated from the Sentinel 2 product (COPERNICUS/S2) at 10 m resolution and averaged over the 2016-2019 period, using Google Earth Engine and the *rgee* R package[66]; average maximum NDVI showed a very strong, linear relationship with the average NDVI estimated during the snow-free season ($R^2 = 0.96$)

Analysis of environmental DNA

DNA analyses were performed as described in ref[55]. Briefly, eDNA was extracted from the dried samples of each plot using the NucleoSpin® Soil Mini Kit (Macherey-Nagel, Germany) with one negative extraction control every ~23 samples. We used a combination of generalist and more specific metabarcoding primers to obtain the broadest representation of soil biodiversity[67] (Supplementary Table S8). First, we used two generalist primer pairs targeting bacteria and eukaryotes to cover the overall variation of biodiversity. Generalist primer pairs were Euka02 (amplifying all the Eukaryota[68]) and Bact02 (Bacteria[6]; note that this marker can yield biased estimates for Archaea diversity[69,70]). We then added five specific primer pairs targeting key taxa within eukaryotes (fungi, vascular plants, and three key animal groups: insects, springtails and earthworms) to maximize the resolution of organisms that have a particularly important role in soil food webs. The specific primer pairs were: Fung02 (amplifying Mycota[71]), Sper01 (seed plants[72]), Inse01 (Insecta[6]), Coll01 (Collembola[73], i.e. springtails), and Olig01 (Oligochaeta[74], i.e. earthworms). All these markers have excellent performance for soil metabarcoding analyses[1,6,75]. Primers were tagged with 8-nucleotide-long tags on the 5′ end to allow bioinformatic discrimination of PCR replicates after sequencing[76]. Extracted DNA was randomized in 96-well plates with extraction controls, blanks, PCR negative and positive controls (total across all plates: 291 blanks, 90 negative and 53 positive controls)[77]. The positive controls comprised genomic DNA of eight bacterial and two fungal strains (see ref[55]), as well as 16 non-tropical plant species belonging to 15 families (Taxaceae, Lamiaceae, Salicaceae, Polygonaceae, Betulaceae, Oleaceae, Pinaceae, Caprifoliaceae, Pinaceae, Aceraceae, Poaceae, Rosaceae, Brassicaceae, Geraniaceae, and Ericaceae). Positive controls allowed us to evaluate cross-contaminations and to assess the performance of amplification and sequencing. The optimal number of amplification cycles was assessed using qPCR[55] (Supplementary Table S1). Amplification was performed as described in ref[55]. Libraries were prepared following the MetaFast protocol[6] and sequenced using the MiSeq (Bact02 and Fung02) or HiSeq 2500 (all others) Illumina platforms (Illumina, San Diego, CA, USA; 2×250 bp for Bact02 and Fung02, and 2×150 bp for the

other markers) at Fasteris (SA, Geneva, Switzerland). The average sequencing depth was ~10,000 reads per PCR replicate for all the markers. Rarefaction curves confirmed that sequencing depth was generally appropriate for all the markers and plots (Supplementary Fig. S1).

Bioinformatic treatment, filtering, clustering and taxonomic assignation were performed using the OBITools software suite[78] and the *sumaclust* program (https://git.metabarcoding.org/obitools/sumaclust/wikis/home) as described in ref[55]. We de-replicated sequences using the *obiuniq* program and discarded bad-quality sequences, sequences shorter or longer than expected and singletons, as their removal is advised to increase the robustness of results[79,80]. We then ran the *obiclean* program to detect potential PCR/sequencing errors, keeping sequences tagged as "heads" in at least one PCR replicate. Sequences were clustered into Molecular Operational Taxonomic Units (MOTUs) using taxon-specific similarity thresholds[81]. Subsequently, we performed additional filtering in R to remove contaminants and spurious sequences[77,82]. We discarded MOTUs observed less than twelve (Euka02), eleven (Coll01), ten (Olig01), eight (Sper01) or five (Inse01, Fung02, Bact02) times overall. These values allow removing ≥ 99.99% of sequences detected in the blanks (i.e. tag-jump errors) of each marker. We also discarded MOTUs detected in just one sample, as they often represent spurious sequences[83], MOTUs detected in <2 PCR replicates of the same sample (possible false positives[84]), and MOTUs detected in >1 extraction or PCR negative control (possible contaminants[77]). Additional details on bioinformatics and filtering are provided by ref[55].

eDNA data enable standardized and cost-effective assessment of soil biodiversity over broad spatial scales[49], but do not produce exactly the same diversity estimates as traditional surveys. Nevertheless, several studies have demonstrated that these two approaches produce highly consistent estimates of diversity[85,86]. Although DNA might in principle persist in the soil for long periods, soil eDNA mostly represents species that are present in the plot during the year of collection[85,87]. Considering the temporal scale of our study (>100 years for all the forelands), the sporadic detection of taxa that were present a few years before sampling is not expected to bias our conclusions.

To further confirm the consistency between eDNA and traditional surveys, we compared the two approaches for a subset of landscapes where both eDNA and traditional surveys were available (plants: 13 forelands from the three continents[47]; insects: 10 forelands from two continents; Supplementary Table S6). For both plants and insects, we found an excellent match between traditional and eDNA data, confirming the robustness of eDNA-based estimates of biodiversity (Extended Data Fig. 7)[47]. See ref[47] for a validation of beta-diversity estimates obtained using eDNA in glacier forelands.

Groups of soil eukaryotes

The Euka02 primers amplify all eukaryotes, but many of them (fungi, plants, insects, springtails and earthworms) were also amplified by more specific primers that target different genomic regions and have better taxonomic resolution[1,6]. Therefore, we first removed from the dataset obtained with the Euka02 primers the MOTUs assigned to fungi, plants, insects, springtails and earthworms[67]. The remaining MOTUs were assigned to: i) protists and ii) "other animals". Because there is not a single accepted definition for protists[88], we defined them as non-fungal unicellular eukaryotes (*sensu* ref[89]). Therefore, based on Simpson[90], protists included: Apicomplexa, Apusomonadidae, Arcellinida, Bacillariophyta, Bicosoecida, Cercozoa, Choanoflagellata, Chrysophyceae, Ciliophora, Cryptophyta, Dinophyceae, Eccrinales, Echinamoebidae, Eustigmatophyceae, Glaucocystophyceae, Hartmannellidae, Hemimastigophora, Heterophryidae, Hyphochytriomycetes, Labyrinthulomycetes, Leptomyxida, Longamoebia, Nucleariidae, Oikomonadaceae, Oomycetes, Perkinsozoa, Picozoa, Plasmodiophorida, Protosteliales, Pterocystidae, Raphidiophryidae, Stemonitida, Synurophyceae, Vampyrellida, , Vannellidae, Variosea, Xanthophyceae. The group of "other animals" contained MOTUs assigned to Arthropoda (excluding Insecta and Collembola), Gastrotricha, Mollusca, Nematoda, Platyhelminthes, Rotifera and Tardigrada.

After filtering, we did not detect vertebrate reads, probably because we used primers amplifying all the eukaryotes. In the soil, vertebrates usually show very low biomass compared to other organisms (e.g. protists, fungi, annelids and arthropods)[91–93]. Therefore, vertebrates provide a limited contribution to the pool of sequences representing all the eukaryotes and remain often undetected in metabarcoding studies that do not use primers specific for this clade[94]. In order to assess the dynamics of vertebrates, future analyses could either use specific primers targeting mammals and / or birds on the extracted DNA, or monitor them using traditional approaches[6,50].

Statistical analyses

General overview

We used the chronosequence approach (space-for-time substitution) to identify the processes underlying ecosystem development in proglacial landscapes. This approach assumes that, within a chronosequence, the time elapsed since glacier retreat provides a maximum age for the ecosystem of a given site, and the comparison of sites deglaciated at different times provides a proxy of ecosystem development over time[95]. We acknowledge that chronosequences are a not a perfect representation of temporal successions, as several additional environmental (e.g. differences in microclimate, topography and substrate) and stochastic (e.g. availability of potential propagules) factors can profoundly affect the development of ecosystems[96–100]. Given the lack of temporally replicated, multi-taxa datasets analyzing biotic colonization of recently deglaciated terrains[101], the chronosequence approach remains the key method of reconstructing long-term (>100 years) trends of ecosystem development[17]. Previous analyses suggest that community patterns retrieved from chronosequences are highly consistent with patterns retrieved from temporally replicated plots[47,50,102]. To validate the conclusions obtained from the space-for-time substitution, we compared the predictions of our models with available data from permanent plots collected by fully independent studies[103,104]. Furthermore, in addition to time since glacier retreat, we considered the possible causal effects of many factors that can influence ecosystem development (microclimate, topography and substrate) and that can potentially confound the chronosequence approach[95]. Additionally, we focused on overall local diversity and compositional dissimilarity, instead of considering species identity, which is particularly affected by local conditions[95].

The aim of this study was to identify the processes determining the development of ecosystems, thus we compared environmental properties within each of the 46 proglacial landscapes. In so doing, we focused on how the properties of each proglacial landscape change along the chronosequence, while considering a large number of landscapes in order to draw general conclusions. This allowed us to understand why plots in the same landscape, but with different ages and habitat properties, have different biodiversity values. The identification of drivers of differences among proglacial landscapes (e.g. why observed richness is larger in one landscape than in another) was beyond the aim of this study, as it would require a macro-ecological and biogeographical perspective.

Variation of environmental properties over time

When using eDNA metabarcoding data, measures of alpha-diversity obtained using Hill's numbers > 0 can provide a more robust estimate of taxonomic richness than the number of MOTUs[82,105]. We thus estimated the local diversity (i.e. alpha-diversity) of the different groups of organisms using the Shannon diversity index (i.e., the exponential of the Shannon entropy or the equivalent of using the Hill's number q = 1). In our analyses, we used the local diversity values estimated using all data obtained after filtering to avoid the loss of rare, but potentially important MOTUs[106]. Nevertheless, the local diversity data were very well correlated to estimates obtained using rarefaction (Supplementary Table S9).

To describe the relationships between time since glacier retreat and environmental properties (soil properties and nutrients, productivity and local diversity), we used generalized linear mixed-effect models (GLMMs) implemented in a Bayesian framework. Environmental properties measured at each plot were the

dependent variables and time was the independent variable. Time was included with both linear and quadratic terms to test the possibility of non-linear relationships. Quadratic terms with confidence intervals that do not overlap zero were considered strong evidence of non-linear relationships. As random effects, we included the identity of the proglacial landscape and site (nested within landscapes), to take into account non-independence of plots within each site and within each landscape.

In all GLMMs, time since glacier retreat was log-transformed and then scaled (mean = 0, SD = 1) to improve normality and facilitate convergence. For values of local diversity, phosphorus, carbon and productivity, we assumed a lognormal error distribution; for pH we assumed a Gaussian distribution; for nitrogen a hurdle Gamma distribution implemented in the *brms* R package[107] due to the large amount of zeroes (i.e., below the detection limit). GLMMs were run on 4 parallel chains of length 50,000 with a burn-in of 5,000 iterations and a thinning rate of 10. Uninformative priors were used as provided in the *brms* package[107]. Analyses were also repeated using weak informative priors. Previous studies suggest a decrease of pH through time and an increase of most nutrients and the alpha-diversity of most taxa[2,24,108], therefore we assumed a weak negative prior for pH (mean = -1, SD = 2) and a weak positive prior for the remaining properties (mean =1, SD = 2). No clear prior knowledge was available for protists, other animals, P and C:P, therefore we used a weak prior with mean = 0 and SD = 2. Models with informative and uninformative priors yielded nearly identical results (Extended Data Table 1). Convergence was assessed for each parameter estimate by visually inspecting the Markov chains and was considered sufficient when $\hat{R} < 1.01$. The absence of spatial autocorrelation was checked for all models by visualizing spline correlograms using the *ncf* R package[109] (Supplementary Fig. S2-S3). The amount of variation explained by GLMMs was assessed using marginal and conditional R^2 (R^2_M and R^2_C, respectively)[110].

We then tested whether models based on the space-for-time substitution provide a reliable estimate of biodiversity changes actually occurring through time. We did a thorough search of the literature to identify studies reporting changes of local diversity occurring through time in permanent plots and found four time series from three glacier forelands (48 temporally replicated observations spanning from 2 to 61 years after glacier retreat)[103,104]. We then used GLMMs to test whether our chronosequence-based models correctly predict the biodiversity changes observed in permanent plots.

Finally, we evaluated whether the development of local diversity follows different patterns in subpolar, temperate and tropical landscapes. Latitudinal groups were defined on the basis of the altitude of the treeline[111,112] (Extended Data S1). Subpolar: areas where the tree line is < 1000 m above sea level (a.s.l.); these are the landscapes higher of 60°N or lower of 45°S. Tropical and sub-tropical: areas where the treeline is >3000 m a.s.l.; these are the landscape between 33°N and 23°S. The broad latitudinal range of this group is justified by the fact that, over a range of about 50° across the equator, treeline position does not change significantly with latitude[111]. Temperate: areas where the treeline is between 1000 and 3000 m a.s.l. In our dataset, this comprises forelands between 42°N and 51°N, and between 34°S and 43°S. Differences in biodiversity development between regions were evaluated by running GLMMs with the interaction between the variables time and latitudinal group.

Drivers of environmental properties and community diversity

We used structural equation models (SEMs)[113] to evaluate the direct and indirect relationships among environmental properties and the local diversity of different taxa. SEMs require the definition of an a-priori model describing the directionality of potential relationships between variables. Soil temperature, time since glacier retreat and TWI were considered to be exogenous variables. Soil pH and nutrients (N and P content) were assumed to be potentially affected by the three exogenous variables. Carbon content showed a very strong relationship with nitrogen content ($R^2_C = 0.9$ in a model relating C to N). Building composite or latent variables in SEMs including random effects is extremely challenging[113] thus, to limit multicollinearity issues, we only included nitrogen into models given the expected importance of soil fertility on community establishment. Models considering C instead

of N yielded equivalent results (Supplementary Table S4). Previous studies have shown that the accumulation of soil organic nitrogen is faster under warmer conditions, therefore we also examined the interaction between temperature and time as a predictor of nitrogen content[24]. The diversity of plant communities was considered to be possibly affected by microclimate, time since glacier retreat and TWI. NDVI was affected by microclimate, time, TWI, soil features and plant diversity. We assumed that plant diversity drives NDVI because our study systems are at the early stages of successions, thus the colonization by plant species is the necessary prerequisite to increase productivity. Finally, the diversity of the remaining soil organisms was potentially affected by temperature, time, TWI, pH and plant diversity. For the diversity of animals and protists, we assumed a potential causal effect of plant diversity, whereas for bacteria and fungi we assumed covariation with the richness of plants, given their frequent symbiotic relationships, for example in nitrogen fixation[43,46,114,115]. Preliminary models suggested a hump-shaped relationship between time and pH, and between pH and the diversity of bacteria, fungi, insects and other animals, with highest diversity values at neutral pH. Therefore, for this variable, we also considered a quadratic term.

The model also included covariation of diversity among all taxa, and covariation among the three soil properties. Piecewise SEMs can integrate random effects into models and thus are particularly appropriate for our dataset, which shows a very complex structure (proglacial landscapes; sites within those landscapes; plots within a site)[113]. Identities of proglacial landscape and site were thus included as random effects in the models. Before running SEMs, we log-transformed local diversity values (q = 1), time, TWI, NDVI, P and N content to improve normality and reduce skewness. All variables were then scaled at mean = 0 and SD = 1. This analysis was limited to the plots where data on both soil features (pH, N and P) and biodiversity were complete (total: 793 pots from 32 landscapes). Given the observed non-linear relationships between time and some ecosystem features (e.g. species diversity of several taxa and some soil features; Fig. 1), we also considered preliminary SEM models assuming a non-linear (quadratic) relationship effect of time. However, these models always showed much higher Bayesian Information Criterion (BIC) values than the simpler models without the quadratic effect of time and thus were discarded. The difference between the SEM models and the GLMMs that analyzed variation of environmental properties through time (Fig. 1, which often showed non-linear effects) probably arose because the GLMMs only included the effect of time, while SEM models also considered the effects of additional variables. For example, if only time is considered as an independent variable, earthworm richness showed a positive quadratic term (Fig. 1), but this quadratic term disappeared when additional variables (soil nutrients, productivity and plant diversity), which also are related to time, were considered. We also tested further variables potentially affecting microclimate (e.g. annual solar radiation and aspect), but none of them significantly improved models (Supplementary Table S10).

Three alternative relationships are possible among diversity of organisms, nutrient availability and productivity: 1) nutrient-led: growing nutrient availability/productivity favors the colonization of soil organisms; 2) biodiversity-led: the development of soil biodiversity fosters an increase of nutrients and productivity; and 3) covariation: instead of a direct causal relationship, there is covariation between the diversity of organisms and nutrient availability/productivity (see Extended Data Figs. 4-5 for detailed representations of the alternative models). We thus used the Bayesian Information Criterion (BIC) to assess the relative support of these models. BIC is an information criterion that can be used to compare models on the basis of their likelihood and number of parameters, and is able to identify the correct SEM structure, particularly if sample size is large[116–118]. Lower BIC values indicate a better supported model.

SEMs were fitted with two alternative approaches that allow the incorporation of clustered data: piecewise structural equation models[113] and lavaan.survey[119]. Piecewise SEMs have high statistical power and provide robust estimates of model parameters[113,117], thus we used them to calculate coefficients, standardized coefficients and effect sizes. lavaan.survey was used to calculate BIC on the basis of the log-likelihood of the multivariate

distribution imposed by the causal structure[119–121]. Fisher's C was used to test the goodness-of-fit of piecewise SEMs[113], while scaled chi-square, root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) were used to test the performance of models fitted using lavaan.survey[122]. The residuals of the best SEM did not show spatial autocorrelation significantly higher than expected by chance (analysis of Moran's I: $P > 0.1$ for all the taxa; Supplementary Table S11), suggesting that autocorrelation of nearby plots was not an issue[123,124]. Spatial autocorrelation analyses were performed using the *EcoGenetics* R package[125].

In order to test whether time, habitat features and biotic interactions play a comparable role in tropical, temperate and subpolar landscapes, we ran separate piecewise SEMs for forelands belonging to each climatic region. From each of these geographically-restricted models, we calculated the absolute value of standardized coefficients representing the direct effects of habitat, time and biotic relationships[126]. Standardized coefficients representing the same process were compared across regions and with the SEM built using all the data.

Finally, in order to evaluate the relative importance of time, habitat features and biotic interactions, we removed from the best BIC model the relationships representing: i) the direct effects of time on local diversity; ii) the direct effects of habitat features (microclimate and soil features) on local diversity; and iii) the covariations between taxa. We iteratively removed one relationship at the time and calculated the resulting change in BIC. The average support of time, habitat and biotic interactions was calculated as the average change in BIC, across all the paths representing each process.

Drivers of community composition

We analyzed differences in community composition considering four taxonomic groups: bacteria, fungi, plants and animals. For each group, sites with zero MOTUs were removed before calculating dissimilarities (Supplementary Table S12). We pooled all the animals (insects, springtails, earthworms and other animals) to increase the number of available pairwise comparisons. Only plots with complete soil features (pH, N, C and P) and with at least one MOTU of bacteria, fungi, plants and animals were retained (total: 5741 pairs of plots from 32 landscapes). Protists were not considered here to reduce the number of plots excluded because of zero MOTUs for any taxon (Supplementary Table S12). We used the Jaccard index to calculate pairwise dissimilarities in community composition (i.e. beta-diversity) between plots belonging to the same proglacial landscape. Previous analyses demonstrated that, in proglacial landscapes, community dissimilarities calculated using eDNA data are highly consistent with dissimilarities calculated using traditional (morphological) approaches[47].

We used SEMs to evaluate how the compositional dissimilarities between plots of the same landscape are related to differences in time and habitat features. For each pairwise comparison, we calculated geographic distance, as well as the absolute differences in time, pH and primary productivity (NDVI). Microclimate dissimilarities and differences in soil fertility were calculated as Euclidean distances using: i) soil temperature and TWI and ii) soil N, C and P, respectively. All variables except coordinates were scaled to mean = 0 and SD = 1 before calculating distances. All distance matrices were log-transformed to improve normality and reduce skewness. The overall structure of SEMs and the causal relationships were similar to the ones used for the analysis of local diversity; we assumed covariation between community composition and nutrient availability/productivity, as this was the most supported model by the analysis of community diversity. Nearby plots are generally more similar than expected by chance, thus we included geographic distance between plots as an additional extrinsic variable, related to all the environmental properties. Glacier identity was introduced as a random intercept to account for potentially different baseline levels of dissimilarity across glaciers.

The analysis of community composition was based on dissimilarity matrices. The non-independence of distance values involving the same plots hampers the use of standard statistics for calculating *P*-values, thus we estimated significance by the permutation of regression parameters[127,128]. Such tests are usually performed by permuting rows and columns of the response matrix. This allows keeping the distance structures fixed and corresponds to a

name-shuffling approach (multiple regression on distance matrices framework)[127,129]. For SEMs, additional complexity is introduced by having multiple matrices that act as both responses and predictors in different portions of the model. Consequently, we generated a permutation scheme where each response matrix was permuted one by one within landscape (9,999 permutations), while keeping fixed the remaining model structure. We first retrieved correlation coefficients (r) and pseudo-t statistics $\left(\frac{b}{\sqrt{1-R_c^2}} \right)$ of the observed model for testing the significance of covariations and causal relationships, respectively, and then stored the same coefficients, at each iteration, for the permuted matrix[127]. P-values were calculated for each response matrix using a two-tailed test, based on the frequency of permuted coefficients (r for covariations, pseudo-t for causal relationships), with absolute values larger than the absolute observed coefficient[127,129]. Also for community composition, standardized coefficients of geographically-restricted models were used to assess the consistency of models across regions of the world. Furthermore, we used the drop in performance (BIC) of models after iteratively removing one relationship at the time (calculated in lavaan.survey) in order to assess the support of candidate drivers.

DATA AVAILABILITY

Raw sequence data and filtered sequence data are available at https://doi.org/10.5281/zenodo.6620359 and https://doi.org/10.5281/zenodo.10423968 digital repositories, respectively.

CODE AVAILABILITY

All codes used are available at https://doi.org/10.5281/zenodo.10423968

AUTHOR CONTRIBUTIONS

G.F.F. conceived the work with the help of W.T., P.T. and J.P.; G.F.F., S.M., A.G., A.B, R.A., M.C., F.A., R.S.A., P.A., P.A.G., S.C.F., J.L.C.L, P.C., M.C.S., J.C., J.A.C.R., C.C., R.C.E., O.D., P.D., A.E., S.E., A.F., L.G., F.G., M.G, S.H., R.K., N.K., R.I.M., M.A.M.M., G.P., F.P., A.R., K.S., L.T., N.U., Y.Y., V.Z., A.Zi., G.A.D. and J.P. planned data collection and performed sampling; A.G., A.B., C.C., L.G., A.P, A.Ze. and G.F.F. performed laboratory analyses; A.C., S.M., A.G., I.C., D.F., W.T. and G.F.F. contributed to data preparation and statistical analyses; A.C. and G.F.F. prepared the first draft of the manuscript with subsequent contribution of all the authors.

COMPETING INTERESTS

The authors declare that they have no conflict of interest.

MATERIALS & CORRESPONDENCE

Correspondence and requests for materials should be addressed to A.C. alexis.carteron@purpan.fr, S.M. silvio.marta@hotmail.it, and G.F.F francesco.ficetola@unimi.it.

METHODS REFERENCES

52. Zemp, M. *et al.* Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016. *Nature* **568**, 382–386 (2019).

53. Marta, S. *et al.* The Retreat of Mountain Glaciers since the Little Ice Age: A Spatially Explicit Database. *Data* **6**, 107 (2021).

54. Dickie, I. A. *et al.* Towards robust and repeatable sampling methods in eDNA-based studies. *Mol. Ecol. Resour.* **18**, 940–952 (2018).

55. Guerrieri, A. *et al.* Metabarcoding data reveal vertical multitaxa variation in topsoil communities during the colonization of deglaciated forelands. *Mol. Ecol.* (2023) doi:10.1111/mec.16669.

56. Rime, T. *et al.* Vertical distribution of the soil microbiota along a successional gradient in a glacier forefield. *Mol. Ecol.* **24**, 1091–1108 (2015).

57. Guerrieri, A. *et al.* Effects of soil preservation for biodiversity monitoring using environmental DNA. *Mol. Ecol.* **30**, 3313–3325 (2021).

58. Bray, R. H. & Kurtz, L. T. Determination of Total Organic and Available Forms of Phosphorus in Soils. *Soil Sci.* **59**, 39–46 (1945).

59. Olsen, S. R. *Estimation of Available Phosphorus in Soils by Extraction with Sodium Bicarbonate*. (Washington, D.C. : U.S. Dept. of Agriculture, 1954).

60. Marta, S. *et al.* Heterogeneous changes of soil microclimate in high mountains and glacier forelands. *Nature Communications* (2023) doi:10.21203/rs.3.rs-2017904/v1.

61. Smith, P. & Metcalfe, P. dynatop: An Implementation of Dynamic TOPMODEL Hydrological Model in R. (2022).

62. Paruelo, J. M., Epstein, H. E., Lauenroth, W. K. & Burke, I. C. ANPP Estimates from NDVI for the Central Grassland Region of the United States. *Ecology* **78**, 953–958 (1997).

63. Rumpf, S. B. *et al.* From white to green: Snow cover loss and increased vegetation productivity in the European Alps. *Science* **376**, 1119–1122 (2022).

64. Lillesand, T., Kiefer, R. W. & Chipman, J. *Remote Sensing and Image Interpretation, 7th Edition*. (Wiley, 2015).

65. Liu, Y. *et al.* Evaluation of consistency among three NDVI products applied to High Mountain Asia in 2000–2015. *Remote Sensing of Environment* **269**, 112821 (2022).

66. Aybar, C. *et al.* rgee: R Bindings for Calling the 'Earth Engine' API. (2022).

67. Ficetola, G. F. & Taberlet, P. Towards exhaustive community ecology via DNA metabarcoding. *Mol. Ecol.* doi:10.1111/mec.16881.

68. Guardiola, M. *et al.* Deep-Sea, Deep-Sequencing: Metabarcoding Extracellular DNA from Sediments of Marine Canyons. *PLOS ONE* **10**, e0139633 (2015).

69. Moll, J. & Hoppe, B. Evaluation of primers for the detection of deadwood-inhabiting archaea via amplicon sequencing. *PeerJ* **10**, e14567 (2022).

70. Hathaway, J. J. M., Moser, D. P., Blank, J. G. & Northup, D. E. A Comparison of Primers in 16S rRNA Gene Surveys of Bacteria and Archaea from Volcanic Caves. *Geomicrobiology Journal* **38**, 741–754 (2021).

71. Epp, L. S. *et al.* New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol. Ecol.* **21**, 1821–1833 (2012).

72. Taberlet, P. *et al.* Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucl. Acids Res.* **35**, e14–e14 (2007).

73. Janssen, P. *et al.* 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**, 306–318 (2018).

74. Bienert, F. *et al.* Tracking earthworm communities from soil DNA. *Mol. Ecol.* **21**, 2017–2030 (2012).

75. Lunghi, E. *et al.* Environmental DNA of insects and springtails from caves reveals complex processes of eDNA transfer in soils. *Sci. Total Environ.* **826**, 154022 (2022).

76. Coissac, E. OligoTag: a program for designing sets of tags for next-generation sequencing of multiplexed samples. *Methods Mol. Biol.* **888**, 13–31 (2012).

77. Zinger, L. *et al.* DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Mol. Ecol.* **28**, 1857–1862 (2019).

78. Boyer, F. *et al.* obitools: a unix-inspired software package for DNA metabarcoding. *Mol. Ecol. Resour.* **16**, 176–182 (2016).

79. Brown, S. P. *et al.* Scraping the bottom of the barrel: are rare high throughput sequences artifacts? *Fungal Ecology* **13**, 221–225 (2015).

80. Alberdi, A., Aizpurua, O., Gilbert, M. T. P. & Bohmann, K. Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution* **9**, 134–147 (2018).

81. Bonin, A., Guerrieri, A. & Ficetola, G. F. Optimal sequence similarity thresholds for clustering of molecular operational taxonomic units in DNA metabarcoding studies. *Mol. Ecol. Resour.* **23**, 368–381 (2023).

82. Calderón-Sanou, I., Münkemüller, T., Boyer, F., Zinger, L. & Thuiller, W. From environmental DNA sequences to ecological conclusions: How strong is the influence of methodological choices? *J. Biogeogr.* **47**, 193–206 (2020).

83. Bálint, M. *et al.* Millions of reads, thousands of taxa: microbial community structure and associations analyzed via marker genes. *FEMS Microbiol Rev* **40**, 686–700 (2016).

84. Ficetola, G. F. *et al.* Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Mol. Ecol. Resour.* **15**, 543–556 (2015).

85. Ariza, M. *et al.* Plant biodiversity assessment through soil eDNA reflects temporal and local diversity. *Methods in Ecology and Evolution* **14**, 415–430 (2023).

86. Pansu, J. *et al.* Long-lasting modification of soil fungal diversity associated with the introduction of rabbits to a remote sub-Antarctic archipelago. *Biology Letters* **11**, 20150408 (2015).

87. Foucher, A. *et al.* Persistence of environmental DNA in cultivated soils: implication of this memory effect for reconstructing the dynamics of land use and cover changes. *Scientific Reports* **10**, 10502 (2020).

88. O'Malley, M. A., Simpson, A. G. B. & Roger, A. J. The other eukaryotes in light of evolutionary protistology. *Biol. Philos.* **28**, 299–330 (2013).

89. Whittaker, R. H. New concepts of kingdoms or organisms. Evolutionary relations are better represented by new classifications than by the traditional two kingdoms. *Science* **163**, 150–160 (1969).

90. Simpson, A. G. B., Slamovits, C. H. & Archibald, J. M. Protist Diversity and Eukaryote Phylogeny. in *Handbook of the Protists* (eds. Archibald, J. M., Simpson, A. G. B. & Slamovits, C. H.) 1–21 (Springer International Publishing, Cham, 2017). doi:10.1007/978-3-319-28149-0_45.

91. Anthony, M. A., Bender, S. F. & van der Heijden, M. G. A. Enumerating soil biodiversity. *Proceedings of the National Academy of Sciences* **120**, e2304663120 (2023).

92. Fierer, N., Strickland, M. S., Liptzin, D., Bradford, M. A. & Cleveland, C. C. Global patterns in belowground communities. *Ecol Lett* **12**, 1238–1249 (2009).

93. Bar-On, Y. M., Phillips, R. & Milo, R. The biomass distribution on Earth. *PNAS* **115**, 6506–6511 (2018).

94. Zinger, L. *et al.* Body size determines soil community assembly in a tropical forest. *Molecular Ecology* **28**, 528–543 (2019).

95. Johnson, E. A. & Miyanishi, K. Testing the assumptions of chronosequences in succession. *Ecol. Lett.* **11**, 419–431 (2008).

96. Makoto, K. & Wilson, S. D. New Multicentury Evidence for Dispersal Limitation during Primary Succession. *Am. Nat.* **187**, 804–811 (2016).

97. Rydgren, K., Halvorsen, R., Töpper, J. P. & Njøs, J. M. Glacier foreland succession and the fading effect of terrain age. *J. Veg. Sci.* **25**, 1367–1380 (2014).

98. Tampucci, D. *et al.* Plant and arthropod colonisation of a glacier foreland in a peripheral mountain range. *Biodiversity* **16**, 213–223 (2015).

99. Vater, A. E. & Matthews, J. A. Succession of pitfall-trapped insects and arachnids on eight Norwegian glacier forelands along an altitudinal gradient: Patterns and models. *The Holocene* **25**, 108–129 (2015).

100. Damgaard, C. A Critique of the Space-for-Time Substitution Practice in Community Ecology. *Trends in Ecology & Evolution* **34**, 416–421 (2019).

101. Smith, J. *et al.* BioDeepTime: A database of biodiversity time series for modern and fossil assemblages. *Global Ecology and Biogeography* **32**, 1680–1689 (2023).

102. Foster, B. L. & Tilman, D. Dynamic and static views of succession: Testing the descriptive power of the chronosequence approach. *Plant Ecol.* **146**, 1–10 (2000).

103. Erschbamer, B., Niederfriniger Schlag, R., Carnicero, P. & Kaufmann, R. Long-term monitoring confirms limitations of recruitment and facilitation and reveals unexpected changes of the successional pathways in a glacier foreland of the Central Austrian Alps. *Plant Ecol* **224**, 373–386 (2023).

104. Fickert, T. & Grüninger, F. High-speed colonization of bare ground—Permanent plot studies on primary succession of plants in recently deglaciated glacier forelands. *Land Degradation & Development* **29**, 2668–2680 (2018).

105. Mächler, E., Walser, J.-C. & Altermatt, F. Decision-making and best practices for taxonomy-free environmental DNA metabarcoding in biomonitoring using Hill numbers. *Mol. Ecol.* **30**, 3326–3339 (2021).

106. McMurdie, P. J. & Holmes, S. Waste not, want not: Why rarefying microbiome data is inadmissible. *PLOS Computational Biology* **10**, e1003531 (2014).

107. Bürkner, P.-C. brms: An R package for bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28 (2017).

108. Ren, Z. & Gao, H. Abundant and rare soil fungi exhibit distinct succession patterns in the forefield of Dongkemadi glacier on the central Qinghai-Tibet Plateau. *Science of The Total Environment* **828**, 154563 (2022).

109. Bjornstad, O. N. & Cai, J. ncf: Spatial Covariance Functions. (2022).

110. Nakagawa, S. & Schielzeth, H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4**, 133–142 (2013).

111. Körner, C. *Alpine Plant Life: Functional Plant Ecology of High Mountain Ecosystems*. (Springer Nature Switzerland AG, Cham, 2021).

112. Paulsen, J. & Körner, C. A climate-based model to predict potential treeline position around the globe. *Alpine Botany* **124**, 1–12 (2014).

113. Lefcheck, J. S. piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods Ecol. Evol.* **7**, 573–579 (2016).

114. Delavaux, C. S., Ramos, R. J., Sturmer, S. L. & Bever, J. D. Environmental identification of arbuscular mycorrhizal fungi using the LSU rDNA gene region: an expanded database and improved pipeline. *Mycorrhiza* **32**, 145–153 (2022).

115. Delavaux, C. S. *et al.* Mycorrhizal types influence island biogeography of plants. *Commun. Biol.* **4**, 1–8 (2021).

116. Bollen, K. A., Harden, J. J., Ray, S. & Zavisca, J. BIC and alternative Bayesian information criteria in the selection of structural equation models. *Struct. Equ. Modeling* **21**, 1–19 (2014).

117. Hertzog, L. R. How robust are Structural Equation Models to model miss-specification? A simulation study. Preprint at https://doi.org/10.48550/arXiv.1803.06186 (2019).

118. Lin, L.-C., Huang, P.-H. & Weng, L.-J. Selecting Path Models in SEM: A Comparison of Model Selection Criteria. *Struct. Equ. Modeling* **24**, 855–869 (2017).
119. Oberski, D. lavaan.survey: An R Package for Complex Survey Analysis of Structural Equation Models. *Journal of Statistical Software* **57**, 1–27 (2014).
120. Shipley, B. & Douma, J. C. Generalized AIC and chi-squared statistics for path models consistent with directed acyclic graphs. *Ecology* **101**, e02960 (2020).
121. Douma, J. C. & Shipley, B. Testing Model Fit in Path Models with Dependent Errors Given Non-Normality, Non-Linearity and Hierarchical Data. *Structural Equation Modeling: A Multidisciplinary Journal* **30**, 222–233 (2023).
122. Westland, J. C. *Structural Equation Models: From Paths to Networks*. (Springer, Cham, Switzerland, 2020).
123. Dormann, C. F. *et al.* Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**, 609–628 (2007).
124. Lichstein, J., Simons, T., Shriner, S. & Franzreb, K. Spatial Autocorrelation and Autoregressive Models in Ecology. *Ecological Monographs* **72**, 445–463 (2002).
125. Roser, L. G., Ferreyra, L. I., Saidman, B. O. & Vilardi, J. C. EcoGenetics: An R package for the management and exploratory analysis of spatial data in landscape genetics. *Mol. Ecol. Resour.* **17**, e241–e250 (2017).
126. Shipley, B. *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference with R*. (Cambridge University Press, Cambridge, 2016). doi:10.1017/CBO9781139979573.
127. Legendre, P., Lapointe, F.-J. & Casgrain, P. Modeling Brain Evolution from Behavior: A Permutational Regression Approach. *Evolution* **48**, 1487–1499 (1994).
128. Martinez-Almoyna, C. *et al.* Multi-trophic β-diversity mediates the effect of environmental gradients on the turnover of multiple ecosystem functions. *Funct. Ecol.* **33**, 2053–2064 (2019).
129. Lichstein, J. W. Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecol.* **188**, 117–131 (2007).
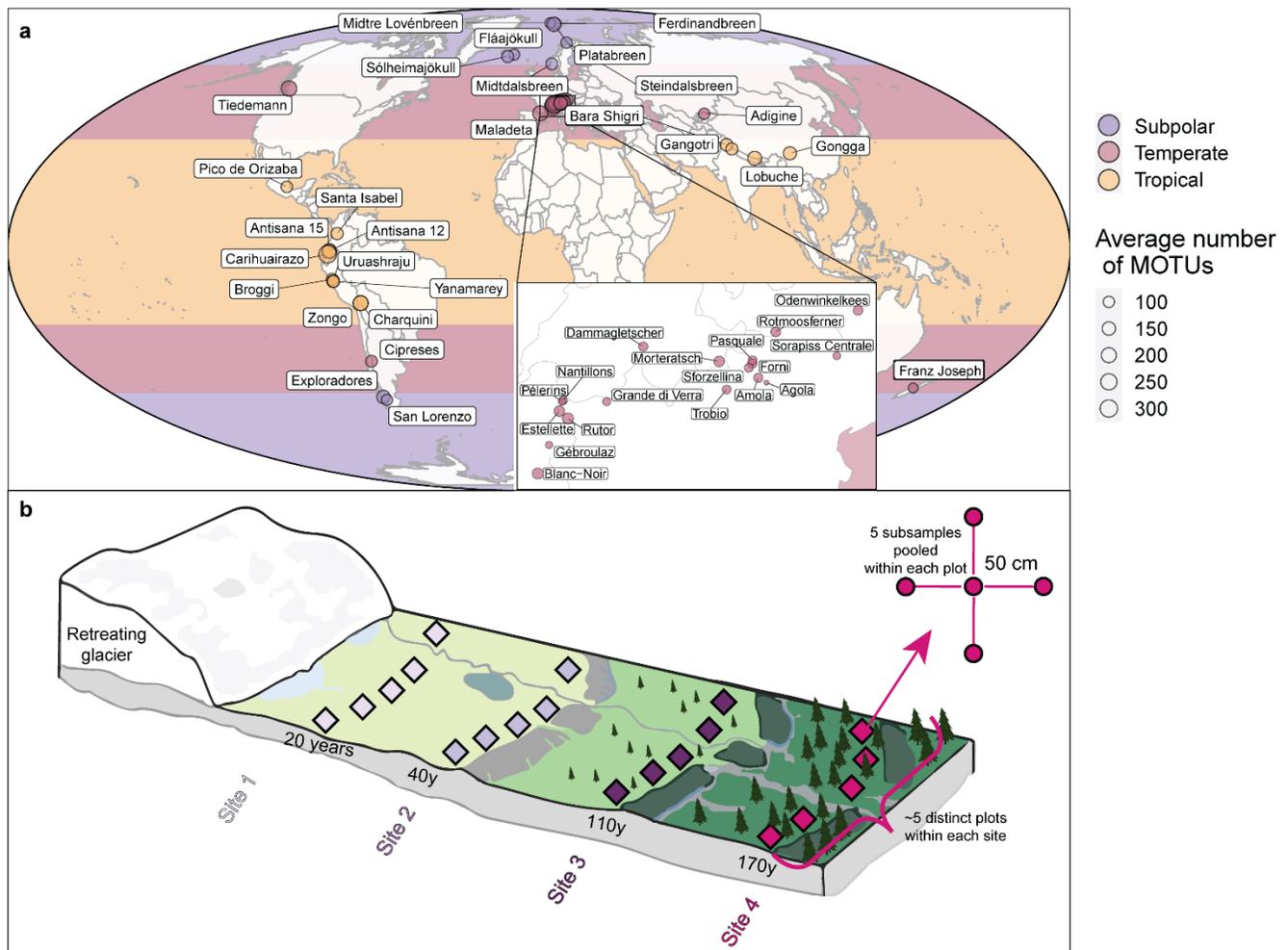
Extended data Fig. 1. a) Global distribution of the 46 analyzed proglacial landscapes. The size of symbols is proportional to the average number of detected MOTUs per plot. The inset shows the landscapes in the European Alps. b) Sampling design. For each proglacial landscape, we selected a number of sites corresponding to the lines representing the position of the glacier forefront at a given date (following ref[53]; four sites are represented in this example). For each site, we established five regularly spaced plots (diamonds; distance between plots: 20 m); at each plot, we collected five soil subsamples within a 1-m radius and pooled subsamples together, resulting in a ~200 g composite sample per plot (total: 46 landscapes; 256 sites; 1256 plots analyzed separately).
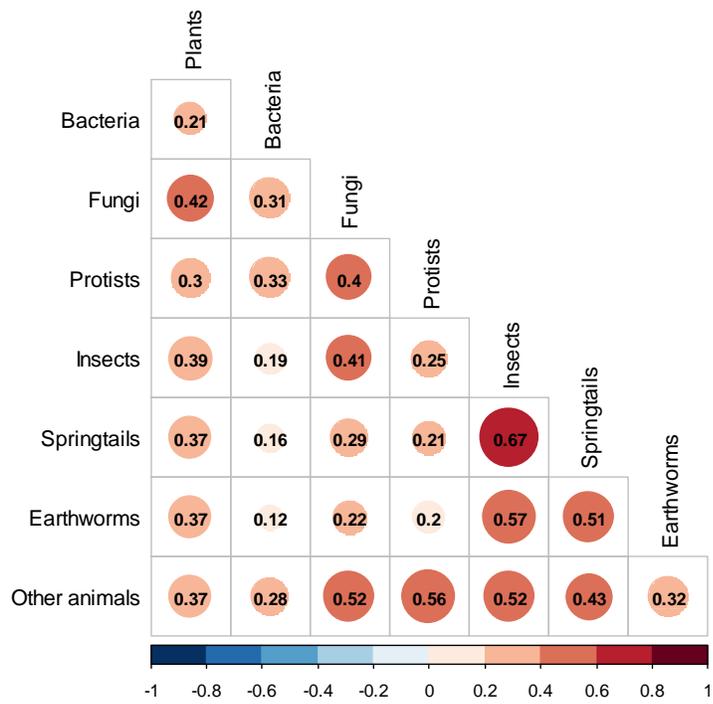
Extended data Fig. 2. Pearson's correlations between the alpha-diversity of eight taxonomic groups of organisms within proglacial landscapes. Larger dots and more intense colors indicate stronger correlations. $N = 1,251$ plots.
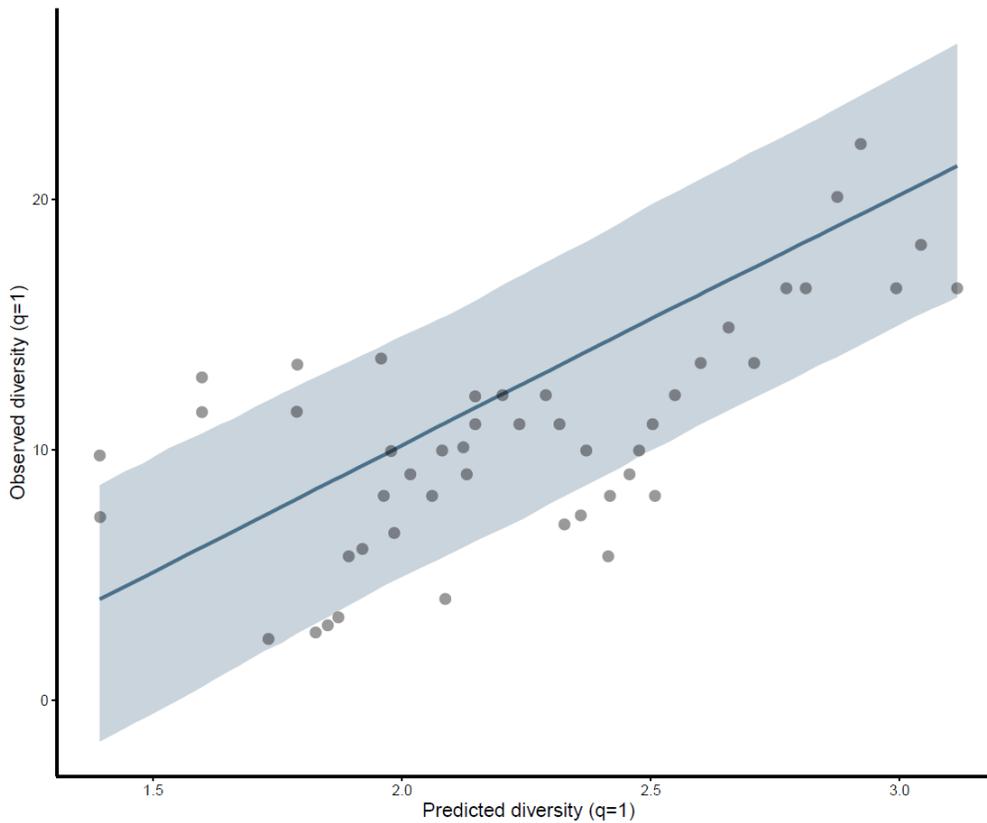
Extended Data Fig. 3. Relationship between diversity values predicted by chronosequence-based models, and diversity values observed in permanent plots surveyed in multiple years by independent studies[103,104]. The bold line indicates values predicted by a mixed model relating observed and predicted diversity (marginal $R^2 = 0.43$, conditional $R^2 = 0.63$), the shaded area represents the 95% credible intervals. Diversity was measured using Hill's number $q = 1$. Differences in absolute values are related to methodological differences between our sampling approach and independent data used for validation. For instance, ref[103] provided total diversity across 10 plots spread over 150 m^2 on each site, ref[3] calculated diversity across 11-13 plots per site, while our diversity predictions refer to one plot per site.
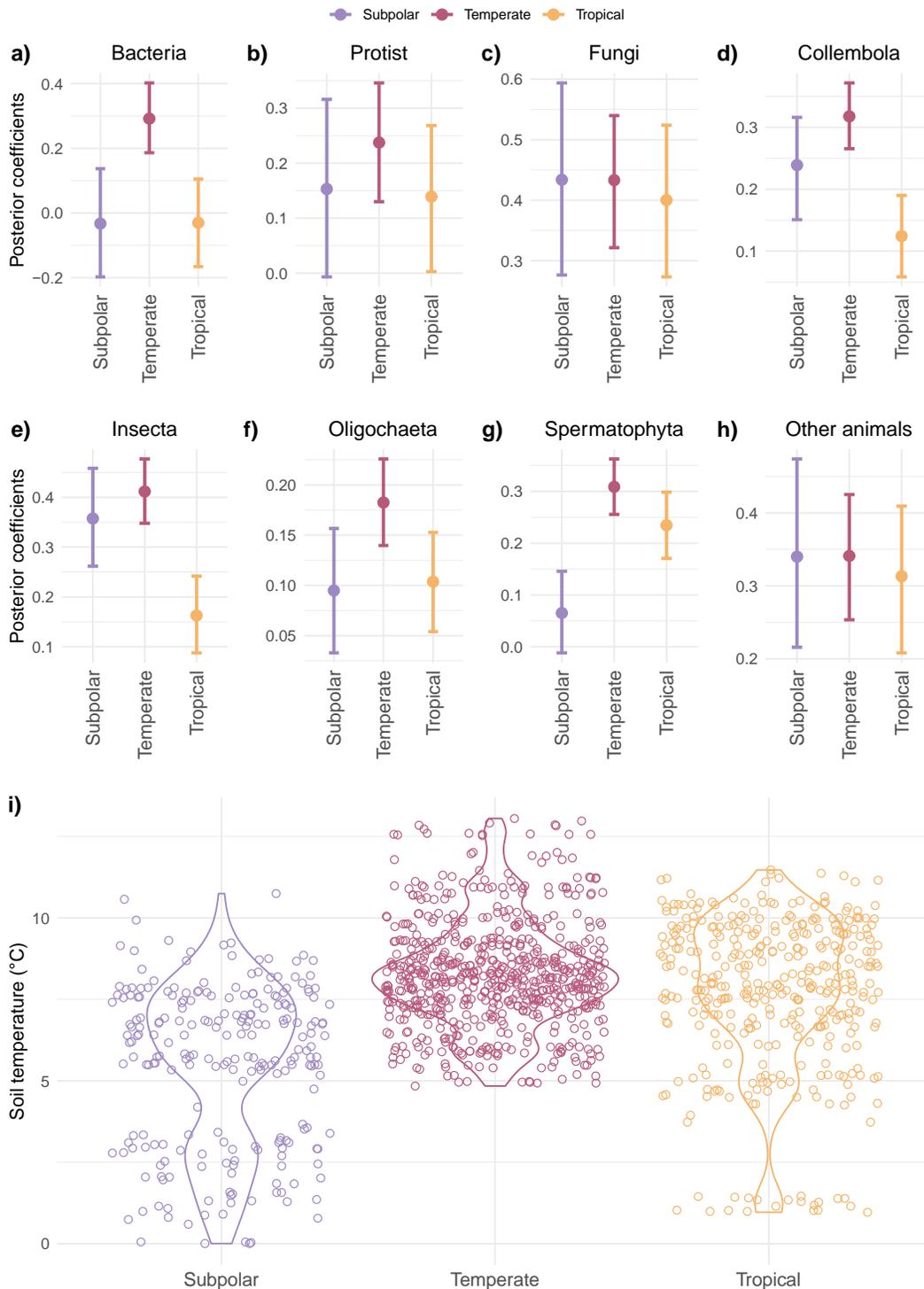
Extended Data Fig. 4. a-h): Differences in colonization rate of eight groups of organisms between subpolar, temperate and tropical landscapes. The plots represent the relationship between age and diversity of the groups estimated by mixed models (Table S2); error bars represent 95% credible intervals. g): Average soil temperature during the growing season in the 1251 analyzed plots.
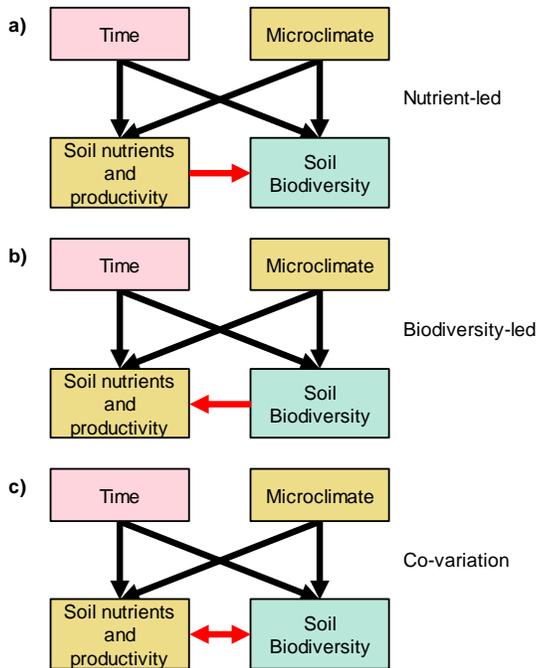
Extended Data Fig. 5. Conceptual models representing how time and microclimate can drive soil chemistry and (alpha or beta) biodiversity changes in ecological succession following glacier retreat. To determine the relationships between soil nutrients and biodiversity, we tested three conceptual causal models assessing three potential causal structures, either (A) soil nutrients and ecosystem productivity shape biodiversity or (B) biodiversity shapes soil nutrients and productivity or (C) soil nutrients, productivity and biodiversity co-vary. "Soil biodiversity" indicates the biodiversity of all the organisms beside plants (i.e. bacteria, fungi, protists and animals). The detailed structure of models, including the relationships between nutrients, plants and productivity, is shown in Fig. 2.
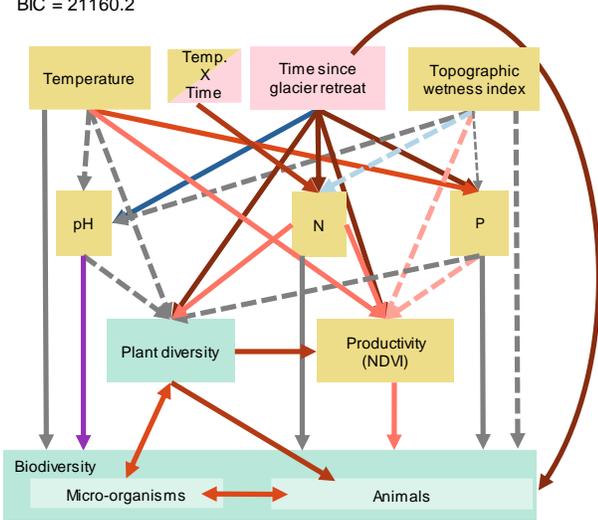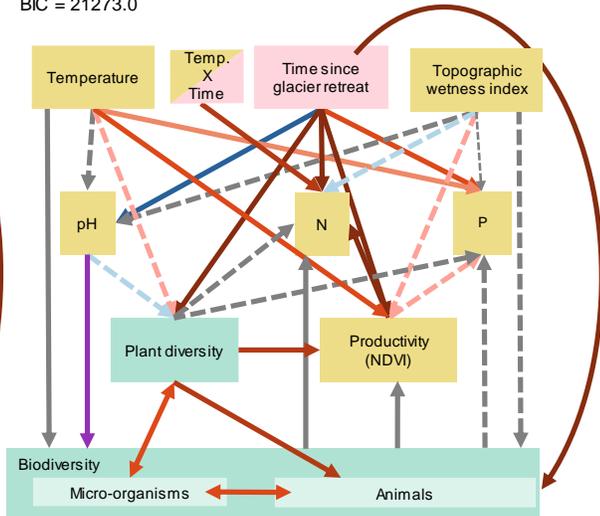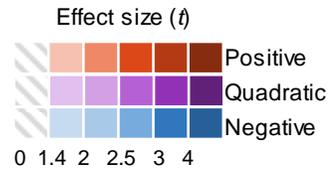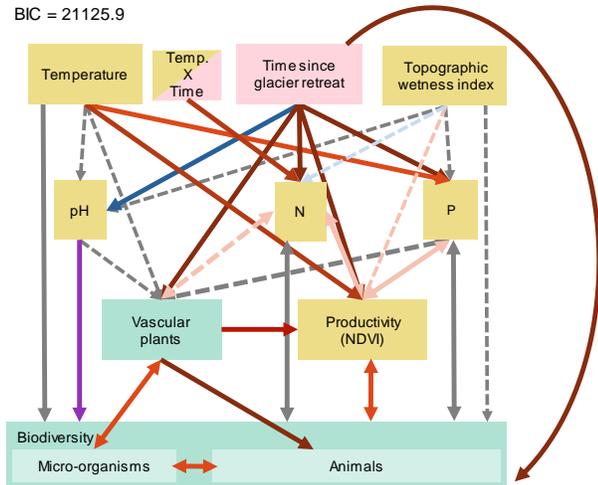
Extended Data Fig. 6. Alternative structural equation models, assuming different relationships between alpha-diversity, soil nutrients and ecosystem productivity. The color and the width of paths is proportional to effect size; dashed lines indicate non-significant relationships. Co-variations between soil features, and between the biodiversity of different taxonomic groups are not shown. In all the models, $N = 793$ plots.
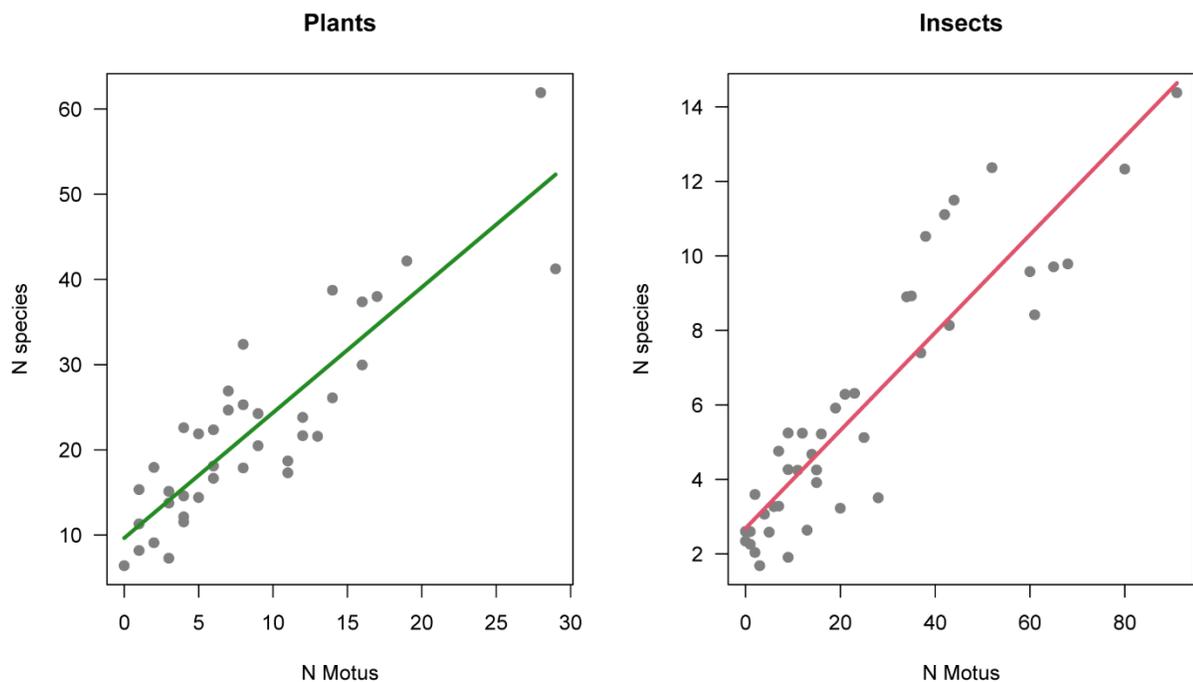
Extended Data Fig. 7. Relationship between the number of Molecular Operational Taxonomic Units (MOTUs) detected by environmental DNA (eDNA) and the number of species recorded in traditional inventories for plants and insects. For plants, N = 38 sites from 10 forelands; for insects, N = 44 sites from 13 forelands. We show the partial regression plot of linear mixed models accounting for glacier identity (plants: $R^2_C = 0.86$; insects: $R^2_C = 0.83$). For eDNA we used the number of MOTUs detected on soil samples across all plots in a site; for traditional detections we used the number of taxa identified at species and genus level. See Table S6 for the sources of traditional data.

Extended Data Table 1. Results of Bayesian GLMMs assessing the evolution of ecosystem attributes and time since glacier retreat as independent variable with conditional and marginal R^2. Estimates with 95% credible intervals. Parameters with 95% credible intervals (CI) non-overlapping zero are in bold.

Dependent variable	N	Independent Variable	B	CI 2.5%	CI 97.5%	Marginal R^2	Conditional R^2	B with informative priors[a]
Phosphorus concentr.	793	Time	**0.55**	**0.40**	**0.70**	0.061	0.636	**0.57**
		Time^2	**0.20**	**0.11**	**0.29**			
Nitrogen concentr.	793	Time	**0.59**	**0.44**	**0.73**	0.015	0.251	**0.59**
		Time2	**0.15**	**0.07**	**0.23**			
Carbon concentr.	793	Time	**1.02**	**0.88**	**1.17**	0.058	0.606	**1.03**
		Time^2	**0.16**	**0.07**	**0.26**			
pH[b]	793	Time	**-0.31**	**-0.40**	**-0.22**	0.048	0.828	**-0.31**
		Time^2	**-0.10**	**-0.15**	**-0.05**			
Productivity (NDVI)	1251	Time	**0.45**	**0.40**	**0.51**	0.193	0.926	**0.45**
		Time^2	**0.05**	**0.02**	**0.09**			
C/N	387	Time	**0.06**	**0.00**	**0.13**	0.006	0.689	**0.06**
		Time^2	0.01	-0.06	0.03			
C/P	631	Time	**0.61**	**0.46**	**0.77**	0.011	0.818	**0.61**
		Time^2	0.02	-0.08	0.12			
Bacteria diversity	1251	Time	**0.11**	**0.03**	**0.19**	0.031	0.525	**0.11**
		Time^2	-0.04	-0.09	0.01			
Protist diversity	1251	Time	**0.18**	**0.11**	**0.26**	0.011	0.440	**0.18**
		Time^2	0.02	-0.03	0.06			
Mycota diversity (fungi)	1251	Time	**0.40**	**0.32**	**0.47**	0.187	0.587	**0.40**
		Time^2	**-0.05**	**-0.10**	**-0.01**			
Collembola diversity (springtails)	1251	Time	**0.23**	**0.19**	**0.27**	0.084	0.421	**0.23**
		Time^2	**0.03**	**0.01**	**0.06**			
Insecta diversity (insects)	1251	Time	**0.30**	**0.26**	**0.35**	0.089	0.504	**0.30**
		Time^2	0.00	-0.03	0.04			
Oligochaeta diversity (earthworms)	1251	Time	**0.13**	**0.11**	**0.16**	0.026	0.299	**0.14**
		Time^2	**0.03**	**0.01**	**0.04**			
Spermatophyta diversity (vascular plants)	1251	Time	**0.23**	**0.19**	**0.27**	0.078	0.384	**0.23**
		Time^2	**0.03**	**0.01**	**0.06**			
Diversity of the other animals	1251	Time	**0.32**	**0.26**	**0.37**	0.110	0.477	**0.32**
		Time^2	-0.01	-0.05	0.02			

[a] Priors were based on the results from refs[2,24,108].

[b] For pH, the credible intervals of the quadratic term overlapped zero (95% CI: -0.13 / 0.12) if plots with age since glacier retreat < 10 years were removed from the analysis.

[c] For C:N ratio, we kept plots with N > 0.01% and removed sites from the Dolomites region, as the carbonate bedrock of this region can affect carbon content.

[d] For C:P ratio, we kept plots with P > 0.0005% and removed sites from the Dolomites region.

Extended Data Table 2. Performance of alternative structural equation models explaining the variation of a) community richness and b) dissimilarity; each SEM is fitted using two alternative approaches: piecewise structural equation models (PSEM) and lavaan.survey. See Fig. 2, Fig, 4 and extended Data 6 for the structure of these models. RMSEA: Root mean square error of approximation; SRMR: Standardized root mean square residual; BIC: Bayesian Information Criterion. The best-BIC model of community richness is in bold. Given the high support of the covariation model in the analysis of community diversity, the same structure was used for the model of community dissimilarity. All significance values are two-sided and derived from a Chi-squared distribution without multiple-test adjustments.

	a. *Dependent: community diversity*								
	pSEM				lavaan.survey				
	Fisher's C	d.f.	P	χ^2	d.f.	P	RMSEA	SRMR	BIC
Nutrient-led model	14.9	10	0.136	29.4	5	<0.001	0.078	0.028	21160.23
Biodiversity-led model	6.6	4	0.162	33.3	4	<0.001	0.096	0.041	21273.01
Covariation model	**9.5**	**8**	**0.301**	**8.62**	**5**	**0.125**	**0.030**	**0.018**	**21125.87**

	b. *Dependent: community dissimilarity*								
	pSEM				lavaan.survey				
	Fisher's C	d.f.	P	χ^2	d.f.	P	RMSEA	SRMR	BIC
Covariation model	4.225	4	0.376	0.030	1	0.861	0.000	0.001	49732.2