Accuracy of mutual predictions of plant and microbe communities varies along a successional gradient in an alpine glacier forefield

Authors: Xie He[1], Maximilian Hanusch[1], Victoria Ruiz-Hernández[1], Robert R. Junker[1,2*]

[1]Department of Biosciences, University Salzburg, 5020 Salzburg, Austria

[2]Evolutionary Ecology of Plants, Department of Biology, Philipps-University Marburg, 35043 Marburg, Germany

*Corresponding author: telephone: +49 6421 28-22434, email: robert.junker@uni-marburg.de

Summary

- Due to climate warming, recently deglaciated glacier forefields create virtually uninhabited substrates waiting for initial colonization of bacteria, fungi and plants and serve as an ideal ecosystem for studying transformations in community composition and diversity over time and the interactions between taxonomic groups.
- In this study, we investigated the composition and diversity of bacteria, and fungi, plants and environmental factors (pH, temperature, plot age and soil nutrients) along a 1.5km glacier forefield. We used random forest analysis to detect how well the composition and diversity of taxonomic groups and environmental factors can be mutually predicted.
- Community composition and diversity of taxonomic groups predicted each other more accurately than environmental factors predicted the taxonomic groups; within the taxonomic groups bacteria and fungi predicted each other best and the taxa's composition was better predicted than diversity indices. Additionally, accuracy of prediction among taxonomic groups and environmental factors considerably varied along the successional gradient.
- Although our results are no direct indication of interactions between the taxa

investigated and the environmental conditions, the accurate predictions among bacteria, fungi, and plants do provide insights into the concerted community assembly of different taxa in response to changing environments along a successional gradient.

Key words: bacteria, environment, fungi, plants, predictive models, succession.

Introduction

Ecological successions represent a sequence of assembly processes leading to diverse and complex communities. It is widely acknowledged that in primary successions stochastic events dominate the initial community assembly whereas niche-based processes become more important in further developed communities where both the environment and species interactions shape species composition and diversity (Chang & HilleRisLambers, 2016; Wojcik *et al.*, 2021). Climate warming is initiating primary successions in glacier forefields as it is speeding up glacier retreat and providing newly uninhabited substrates for the colonization of organisms forming communities along the chronosequence. Local microclimatic conditions and soil properties, and the tight interactions between plant and belowground microbes are part of the environmental and biotic factors shaping these communities (Zak *et al.*, 2003; Mouhamadou *et al.*, 2013; Darcy *et al.*, 2018; Navratilova *et al.*, 2019; Harrison *et al.*, 2020; Ohler *et al.*, 2020; Davison *et al.*, 2021). The relative contributions of environmental and biotic factors on communities may vary spatially and temporally and may affect different properties of communities (Mitchell *et al.*, 2011; Chen *et al.*, 2017). In the case of microbial communities the relative importance of these factors partly depends on the soil compartment: Microbes colonizing the rhizosphere are directly affected by the sugars, organic acids and amino acids provided by root exudates, whereas microbes colonizing bulk soil are less affected by specific plant individuals and may thus respond more strongly to the environmental conditions (Hartley *et al.*, 2007; Lange *et al.*, 2015; Sanchez-Canizares *et al.*, 2017). Accordingly, both environmental factors such as soil chemical properties and temperature (Cheng *et al.*, 2020; Hermans *et al.*, 2020; Davison

et al., 2021) as well as plant community composition (Chen *et al.*, 2017; Kruger *et al.*, 2017; Reese *et al.*, 2018) were reported to explain variation in soil microbial communities. Likewise, microbes and the environment have been shown to affect plant species composition (Bever *et al.*, 2012; Miller *et al.*, 2020). For instance, recent studies demonstrated that bacterial and fungal communities in soil may either positively or negatively affect plant species and communities, and that these effects are tightly related to the plant functional type or the partnership between the plants and microbial organisms (Teste *et al.*, 2017; Hahl *et al.*, 2020; Heinen *et al.*, 2020). Within soil microbial communities, the various interactions between bacteria and fungi are additionally contributing to community assembly. The interactions between microbes that share a habitat can either be positive (mutualism, synergism, or commensalism), negative (pathogenic, predation, parasitism, antagonism, or competition) or neutral (no observed effects) (Vandenkoornhuyse *et al.*, 2007; Berg *et al.*, 2020). In diverse communities all of these outcomes of pairwise interactions may occur simultaneously, which lead to changes of the organism performance and ecosystem productivity (Wargo & Hogan, 2006; Miransari, 2011). The outcome of pairwise interactions between plant, bacterial, and fungal species is highly context dependent and thus modulated by the presence of other species as well as environmental conditions. For instance, environmental conditions like temperature and soil moisture affect plant and microbes and can regulate plant-microbe associations (Rasmussen *et al.*, 2019; Rudgers *et al.*, 2020; Robroek *et al.*, 2021), and increasing environmental stress alters microbial facilitation of plant germination or biomass production (David *et al.*, 2020).

Successional gradients such as glacier forefields with considerable variation in soil properties and climate conditions are an ideal study system to reveal how the interdependences between taxonomic groups change with environmental conditions. Previous studies have shown different interactive patterns of organisms in different successional stages. For instance, a positive relationship between plant and microbial richness was found in the early succession while not in the late succession (Porazinska *et al.*, 2018), which may be owing to the fact that pioneering plants serve as nutrition

hotspot for microbes in the rhizosphere at the initial sites (Schulz *et al.*, 2013) while in the late succession the accumulated organic matter provide plenty resources and thus the mutual predictability between plants and microbes is reduced (Porazinska *et al.*, 2018). Nonetheless, the interdependencies between the organisms that form complex communities as well as environmental conditions may also leave a signal in community composition and diversity of plants and microbes and thus these properties may be mutually predictable (Horn *et al.*, 2017; Leff *et al.*, 2018).

Machine learning algorithms have been increasingly applied for pattern recognition and predictions using complex ecological data. For instance, random forest analysis was used to explore the links between soil bacterial community composition and environmental factors such as land use management and soil properties and lead to predictive models with high accuracy (Hermans *et al.*, 2020). Furthermore, machine learning models were shown to outperform regression models in trait-matching predictions for understanding interaction networks (Pichler *et al.*, 2019). The high performance of machine learning algorithms and especially random forest is obtained by their ability to model non-linear combinations of numerical and categorial data without complex transformations resulting in estimates of the accuracy of predictions as well as the importance of individual variables in improving predictions (Breiman, 2001; Ghannam & Techtmann, 2021). Thus, random forest is an excellent tool for evaluating the interdependences between various taxonomic groups and environmental factors (Ghannam & Techtmann, 2021; Goodswen *et al.*, 2021).

In recent studies, individual plant variables such as species composition, functional identity, taxonomic, phylogenetic and functional diversity have been used to predict microbial communities (Prober *et al.*, 2015; Dassen *et al.*, 2017; Chen *et al.*, 2018; Leff *et al.*, 2018; Porazinska *et al.*, 2018). In the present study, we consider all the possible variables and evaluate how well the diversity and composition of plants, bacteria, fungi, and environmental factors can predict each other in order to explore the interdependences between the taxonomic groups and environmental factors as well as their changes along a successional gradient in the forefield of the Ödenwinkelkees

glacier in the Austrian Alps (Junker *et al.*, 2020). Here the assembly of multidiverse communities along the glacier forefield chronosequence provides an excellent opportunity to track transformations in community composition and diversity over time. Using random forest analysis, we used plant, bacteria, fungi composition and environmental factors as explanatory variables and bacteria, fungi and plant species composition, plant functional composition, plants phylogenetic, functional diversities and environmental factors as dependent variables to test for interdependencies between these variables. We aim to address the following two questions: 1) How accurately can we predict properties of plant and microbial communities with the composition of the other taxonomic groups as well as environmental factors? 2) Is the accuracy of prediction variable along the successional gradient? Although our approach does not directly indicate interactions and dependencies between taxonomic groups and environmental conditions, it tests the hypothesis that taxonomic groups respond to changing environmental conditions in a concerted way potentially facilitated by tight interaction networks.

Materials and methods

Data collection

Our study site was located at the forefield of the Ödenwinkelkees glacier (Stubachtal valley, Hohe Tauern National Park, Austria; Dynamic Ecological Information Management System – site and dataset registry: https://deims.org/activity/fefd07db-2f16-46eb-8883-f10fbc9d13a3, last access: August 2021) (Junker *et al.*, 2020). In summer 2019 (26 June - 16 September), we established 135 permanent plots within the successional gradient of the glacier forefield. We identified all vascular plant species occurring at the plots (n = 107) and recorded the coverage of plants with a resolution of 0.1%. We measured the plant height, leaf area, leaf weight and calculated the specific leaf area (SLA) for those 48 plant species that occurred in 10 or more plots. For three focus species we phenotyped up to three individuals on every plot where they occurred: *Oxyria digyna* as representative of early succession, *Trifolium badium* as representative

of late succession, and *Campanula scheuchzeri* which occurred all along the successional gradient (for detailed information on the selection of the focus plant species see Junker et. al 2020). For the other $n = 45$ species, up to five individuals per plot were phenotyped on the youngest, the oldest, and the intermediate plot where they occurred (for detailed methods see Junker et al 2020). Additionally, we obtained the functional traits of the plant species from Bioflor database (https://www.ufz.de/biolflor/index.jsp) for 92 species out of 107 plant species occurring in the field. We used nine functional traits which have been shown to be response traits to environmental changes at the community level (Kahmen & Poschlod, 2004; Bernhardt-Römermann *et al.*, 2008; Aguiar *et al.*, 2013; Hintze *et al.*, 2013), including fruit type, leaf anatomy, leaf persistence, life form, life span, pollen vector, strategy type, type of reproduction. We also characterized the soil microbiome (bacteria and fungi) of each of the plots. We sampled soil from each plot at two locations at a depth of 3cm, soil from two locations per plot were pooled to one sample for further analysis. Samples were directly transferred to ZR BashingBeads Lysis tubes containing 750 µL of ZymoBIOMICS lysis solution (Zymo-BIOMICS DNA Miniprep Kit; Zymo Research, Irvine, California, USA). Within 8h after collection of microbial samples, ZR BashingBeads Lysis tubes were sonicated for 7 min to detach microorganisms from the surfaces. Subsequently, all microbial samples were shaken using a ball mill for 9 minutes with a frequency of 30.0 s^{-1}. Microbial DNA was extracted using the ZymoBIOMICS DNA Miniprep Kit following the manufacturer's instructions. Microbiome analysis was performed by Eurofins Genomics (Ebersberg, Germany) using the company's standard procedure. To assign taxonomic information to each OTU, DC-MEGABLAST alignments of cluster representative sequences to the sequence database were performed (Reference database: NCBI_nt (Release 2018-07-07)). Further processing of OTUs and taxonomic assignments was performed using the QIIME software package (version 1.9.1, http://qiime.org/) (Caporaso *et al.*, 2010). Abundances of bacterial and fungal taxonomic units were normalized using lineage-specific copy numbers of the relevant marker genes to improve estimates (Angly *et al.*,

2014). Prior to the statistical analysis of microbial communities, we performed a cumulative sum scaling (CSS) normalization (R package metagenomeSeq v1.28.2) on the count data to account for differences in sequencing depth among samples.

To record the seasonal mean temperature, we buried temperature loggers with a resolution of 0.5 °C (MF1921G iButton, Fuchs Elektronik, Weinheim, Germany) 10 cm north of each plot center, at a depth of 3 cm below ground (Junker *et al.*, 2020; Ohler *et al.*, 2020) during field work in 2019. The thermo loggers were set to start on 13th August 2019 and were stopped on 9th August 2020 with a total of 2048 measurements recorded on 362 days. Seasonal mean temperature was calculated on the basis of the recordings ranging from 13th August to 16th of September 2019 and 26th June to 9th August 2020 representing the period in which the plots were free of permanent snow cover before and after the winter 2019/2020. In 2020 (25 July - 21 August), we took additional soil samples from all plots to measure soil nutrient content (N P, K, Mg) as well as soil pH. Samples were sent to AGROLAB Agrar und Umwelt GmbH (Sarstedt, Germany) for analysis.

Data analysis

To test the predictability of the diversity and composition of each of the taxonomic group by the composition of other taxonomic groups as well as by environmental parameters, we used the machine learning algorithm random forest (R package randomForest). Random forest combines several randomized decision trees and aggregates their predictions by averaging, it can handle multiple input variables (explanatory variables), which are ranked by different levels of importance in predicting the dependent variable (Breiman, 2001; Biau & Scornet, 2016). As explanatory variables we used the community tables of plants, bacteria, and fungi with plots as rows and the abundance of the species or OTUs as columns (Table S1, S2 & S3); meanwhile we used multivariate datasets informing about the environmental conditions of each plot with plots as rows and environmental variables as columns (Table S4). As dependent variables we used univariate variable including plant Shannon, phylogenetic and functional diversity, bacteria Shannon diversity, fungi Shannon

diversity, soil seasonal mean temperature, pH, plot age, soil N, P, K, and Mg as well as principal components of the composition of all the taxonomic groups, resulting in 20 variables in total (Table S5). As random forest analysis can only deal with univariate dependent variables, we used the first two principal component axis (PCA) which carry most information of the composition to refer to plant species composition (15.3% + 11.2%), bacteria composition (6.4% + 4.6%) and fungi composition (4.1% + 3.2%). Plant functional composition matrix was generated based on the plant species composition table and the functional traits table obtained from Bioflor database. For each category of each trait, we calculated the total coverage of species belonging to the category, and this was done for all the 9 traits and all 9 traits were merged to a single table, thus generating the functional composition table with plots name as rows and 50 trait categories as columns, i.e. each categorial traits had two or more categories resulting in a total of 50 categories. (Table S6). Plant functional composition was represented by the first two PCAs, too (59.3% + 12.3%). Plant Shannon diversity was calculated from the compositional dataset using the R package vegan (Dixon, 2003). Plant phylogenetic diversity was calculated using the R package picante (Kembel *et al.*, 2010). We extracted a phylogenetic tree using the R package pez (Pearse *et al.*, 2015) for species existing in our field site from a dated molecular phylogeny tree (32,223 species) for land plants (Zanne *et al.*, 2014). In cases where species were not included in the tree, it was substituted by species from the same genus. Among 107 species existing in our plots, we were able to match and built a tree with 104 species and we used it for the calculation of phylogenetic diversity. We used 'Functional dispersion' calculated from the R package FD (Laliberte & Legendre, 2010) as the index for plant functional diversity. The mean plant height, leaf area, leaf weight and SLA of every species were used for the trait table (Table S7) identically for every plot, and for the community table the species with a low occurring frequency along the successional gradient (not included in the 48 species with traits measured) were ignored in the calculation of functional diversity. For bacteria and fungi, the Shannon diversity was calculated based on the OTU composition after rarefying the data to the minimum

number of reads available in the samples (repeats = 999).

Using all combinations of explanatory and dependent variables, we performed random forest analyses with 10-fold cross validations to quantify the performance of the predictive model. Specifically, for each prediction, 80% of the plots were randomly selected as the training dataset and the remaining 20% of the plots were used as test dataset. The predictive model resulting from the training dataset was applied to the test data and the predicted values of the plots in the test dataset were correlated with the observed values of these plots. This process was repeated for ten times, and then we defined the mean Pearson's *r*-value of ten correlations as 'accuracy of prediction' and used the proportion of significant correlations (p-value < 0.05) out of the 10 correlations as 'significance frequency'. Additional to random forest analysis using all the plots for a global impression on the predictability of dependent variables, we also employed a moving frame approach to detect how the predictabilities change along the successional gradient. With the 135 plots, we grouped every 45 plots into one frame and used the median plot as identifier of the frame. Thus, the first frame included plots 1 to 45, the second 2 to 46, and so forth. This approach led to a set of 91 moving frames whose identifiers ranged from plot 23 to plot 113. Using the same proportion of training and test dataset, for every 45 plots in each frame, data of 36 (80%) randomly selected plots was used as training dataset, and the other 9 (20%) plots were used as test dataset. The accuracy of prediction and significance frequency were calculated for every frame as stated before. We fitted a linear or quadratic regression with the accuracy of prediction of every variable along the successional gradient as independent variable and the frame number as explanatory variable. The model with a higher r^2 value was chosen and the statistically significant relationships were shown as a regression line.

To make a comprehensive comparison of how well every variable is predicted by the other individual group and by the other three groups combined, we did the same random forest predicting procedure for each of the 20 variables using the other three groups together (except for the group that was considered in the dependent variable). We compared for each variable how well they were predicted by every other single group

and by three groups combined using the Tukey Test.

Results

In total we obtained soil bacteria and fungi composition data from n = 127 and 130 plots, respectively; n = 5221 bacteria OTUs and n = 6016 fungi OTUs were detected in all the soil samples. Raw sequences of next-generation 16S rRNA gene amplicon sequencing are available at the NCBI Sequence Read Archive (SRA) under the BioProject accession PRJNA701884 and PRJNA701890. The mean accuracy of prediction of each pair of explanatory variables and dependent variables did usually not strongly differ between the global analysis considering all plots and the mean of the frame-wise analyses, indicating the validity of using the moving frames for random forest predictions. Most of the predictions fit a quadratic regression, indicating a non-monotonic change of the accuracy of prediction along the successional gradient.

Bacterial communities as predictors (Fig. 1 and Fig. 5a) – Bacterial communities (quantitative OTU tables) most accurately predicted the taxonomic composition of fungal communities (PC1 and PC2) followed by plant functional composition. Among the environmental parameters, plot age and pH-value were most accurately predicted by bacterial communities. Note that our results do not imply a direction of effects in the sense that the dependent variable is affected by the explanatory variable. For instance, bacterial communities do not affect the soil temperature but are affected by this environmental parameter. Accuracy of prediction of target variables associated with plant communities mostly decreased with plot age, whereas accuracy of prediction of fungi and environmental target variables remained constant or even increased along the age gradient in most cases.

Fungi communities as predictors (Fig. 2 and Fig. 5b)– Fungal communities (quantitative OTU tables) most accurately predicted the taxonomic composition of bacterial communities (PC1 and PC2) and bacterial Shannon diversity was the variable with the lowest accuracy of prediction. Plot age and pH were also the environmental factors that were most accurately predicted by fungi communities. Similar to bacterial

predictions, accuracy of prediction of target variables associated with plant communities mostly decreased with plot age, whereas accuracy of prediction of bacterial and environmental target variables remained constant or increased along the age gradient in most cases.

Plant communities as predictors (Fig. 3 and Fig. 5c)– Plant communities (quantitative vegetation table) predicted the plot age most accurately, followed by fungi composition (PC1) and bacteria composition (PC1). Plant communities predicted bacteria and fungi Shannon diversities least accurately. The plant predictions of variables concerning bacteria, fungi and some environmental parameters were mostly decreasing with increasing plot age. For environmental variables, the accuracy of prediction for temperature, pH and soil Mg increased and the others were mostly decreasing with plot age.

Environmental factors as predictors (Fig. 4 and Fig. 5d)– Environmental factors (multivariate table of environmental parameters) predicted the fungi composition PC1 and bacteria composition PC1 with the highest accuracy, followed by plant functional diversity and plant species composition PC2. Accuracy of prediction for plant variables were mostly decreasing along the gradient, and for bacteria and fungi they either had the highest accuracy of prediction in the middle age or increase with plot age.

Combined groups as predictors (Fig. 6) – Using all taxonomic groups and environmental variables (except for the group that was considered in the dependent variable) together to predict dependent variables increased the accuracy of prediction for plant functional composition PC1, plant functional composition PC2, fungi composition PC1. For plot age, the accuracy of prediction even decreased, and for other variables especially environmental variables, the accuracy of prediction with all the groups combined did not significantly increase.

Discussion

Our results indicate that the composition and diversity of plant, bacteria, and fungi is - to a certain degree - predictable by the composition of the respective other taxonomic

groups as well as by environmental factors. The accuracy of prediction, however, varied along the successional gradient of the forefield of the Ödenwinkelkees glacier. Overall, the taxonomic groups predicted each other more accurately than environmental factors predicted the taxonomic groups; within the taxonomic groups their composition was better predicted than diversity indices. When using a combined dataset as predictors, only a few variables obtained increased accuracy of prediction compared with using a single group, and most of the variables have no significant difference or even decreased accuracy of prediction. Well performing predictive models may indicate direct interactions between taxa or effects of the environment on taxa. However, statistical associations between taxa may also suggest that both taxa respond similarly to a third taxonomic group or an environmental factor (Blanchet *et al.*, 2020). Thus, while our results do not directly reveal ecological interactions, they do provide insights into the concerted community assembly of different taxa in response to changing environmental factors along a successional gradient.

Variables describing the composition of taxonomic groups (e.g. PC axis of community composition) were mostly more precisely predicted by other taxonomic groups than diversity indices. Particularly, the community composition of bacteria and fungi mutually predicted each other most precisely, which confirms previous studies demonstrating the interdependences between bacteria and fungi (Miransari, 2011; Deveau *et al.*, 2018). Both bacteria and fungi community composition predicted plant functional composition more precisely than species composition and functional, phylogenetic and taxonomic diversity of plants. These results indicate that the plant functional identity has a stronger effect on soil microbial communities than plant species identity and diversities (Dassen *et al.*, 2017). Fungal composition was better predicted by plant composition than bacterial composition, which may reflect the tight interaction between plants and fungi, especially mycorrhiza (Horn *et al.*, 2017; Sweeney *et al.*, 2021). The interactions between plants and microbes are mediated through plant root exudates and litter input (Knelman *et al.*, 2012; Lopez-Angulo *et al.*, 2020). Root exudates vary substantially between different plant species and various

343 microbes utilize the carbon source from plants (Vandenkoornhuyse *et al.*, 2007). In this
344 way, the plant community provides various niches for the microbes and plays an
345 important role shaping microbial communities in the soil (Bever *et al.*, 2012). Likewise,
346 the interplay of facilitative and antagonistic effects determines the direction of
347 feedbacks from soil microbes to plants and maintains the diversity of plant communities
348 (Bever *et al.*, 2012; Teste *et al.*, 2017; Mony *et al.*, 2021). Nevertheless, although it has
349 been reported that plant composition has an effect on microbial richness (Lopez-Angulo
350 *et al.*, 2020), we did not detect a strong accuracy of prediction between plant
351 composition and bacterial or fungal Shannon diversity. This suggests that interactions
352 within taxonomic groups are reducing the accuracy of prediction between the
353 composition and diversity of plants and microbes. For instance, positive or negative
354 effects of individual bacterial strains on plant growth may be changed by the presence
355 of other strains (Raza *et al.*, 2020), which may lead to a hardly predictable complexity
356 of interdependencies and influences.

357 Plot age, soil temperature and soil pH were well predicted by taxonomic groups, and
358 soil nutrients were less well predicted. In contrast, the environmental variables did not
359 accurately predict the composition and diversity of the taxa. As stated above, our
360 approach is not implying a direction of effects, which means that it is more likely that
361 the environmental factors affect the composition and diversity of the taxonomic groups
362 and not *vice versa*. Among all the environmental factors, plot age is the environmental
363 factor best predicted by taxonomic groups, followed by soil temperature and pH. Plants
364 predicted plot age better than bacteria and fungi, and the signal was even blurred when
365 using all the groups together. This indicates that plant communities follow a clear
366 succession with age-specific stages. Microbes may be more responsive to other
367 environmental factors that may act on short term fluctuations such as temperature,
368 which is equally well predicted by the compositions of bacteria, fungi and plants
369 suggesting its common importance in defining the niche of all taxa. Previous studies
370 demonstrated that plants and microbes from different origins may respond to increased
371 temperature variously, thus we may infer that climate change will shift the interactive

372 patterns between species (Rasmussen *et al.*, 2019; Rudgers *et al.*, 2020; Losapio *et al.*,
373 2021). In addition, pH was better predicted by bacteria and fungi than by plants,
374 indicating that pH is affecting soil microbes more than plants, which is in agreement
375 with previous studies illustrating the importance of pH in affecting microbial
376 communities (Knelman *et al.*, 2012; Shen *et al.*, 2020). Soil nutrients such as N, P, K,
377 Mg were more accurately predicted by plants than by microbes suggesting strong
378 feedbacks between soil nutrients and plant communities (Fischer *et al.*, 2019). In
379 summary, we showed that plants, bacteria, and fungi mutually predict each other's
380 diversity and community composition and that environmental parameters are also well-
381 suited predictors for the same biotic dependent variables. This is in line with previous
382 studies demonstrating that plant communities and environmental factors are
383 contributing and explaining different parts of variation in soil microbial communities
384 (Mitchell *et al.*, 2011) and that interactions between plants and microbes can be
385 independent on environmental changes (Sweeney *et al.*, 2021).

386 Accuracy of prediction varied with successional age. For instance, plant taxonomic and
387 functional composition was better predicted by bacteria and fungi at early than late
388 succession. This could be explained by a relatively clear signal of interaction between
389 individual plant species and microbes at early succession while the signal of individual
390 plant species may be diluted at late successional stages where communities become
391 complex (Porazinska *et al.*, 2018). The interactions between plants and microbes are
392 known to be responding to primary successions. For instance, while plant-derived
393 carbon becomes a major source for bacteria after 50 years of succession, these
394 communities utilize ancient carbon in the first decades after deglaciation, which has
395 been demonstrated in the area of our study site (Bardgett *et al.*, 2007). In accordance
396 with this finding, Tscherko *et al.* (2005) found evidence of plants shaping microbial
397 communities in soils older than 43 years in another Austria glacier, the Rotmoosferner.
398 These results suggest a higher accuracy of prediction of microbial communities by
399 plants at later successional stages, which is not fully in line with our findings. In contrast
400 to many other statistical methods, random forest decision trees consider individual

features instead of multivariate representations of the communities. Thus, even though bacterial communities are mainly shaped by abiotic factors and non-plant related carbon sources, random forest is able to select those strains that may be associated with the few plant species colonizing the young plots, which represents a strong signal in the data. In contrast in older plots, when plants provide the major carbon source, the signal of each individual species may be diluted resulting in a poor prediction. Additionally, further carbon sources accumulate such decomposed soil organic matters, which again sustains microbial communities unrelated to plant species diversity and composition. Another reason for our finding may be the reduced variability in plant species composition and diversity between older plots, which is a common finding in primary successions (Ortiz-Alvarez *et al.*, 2018). This could also partly explain the decrease of accuracy of prediction between plants and microbes along the succession as the decreased variation of community composition makes it less sensitive to detect the change of the interacting taxa. Finally, age is not the only factor that is affecting the successional age of plots in glacier forefields, instead allogenic factors may reset successions or at least slow down successional progress in community development (Wojcik et al. 2021). These allogenic factors, such as geomorphic events, accumulate over time and thus may lead to outliers in community composition. If these outlier plots are part of test dataset, they cannot be predicted on models as predictions are only possible in the range of the training dataset.

Our results demonstrate the concerted development of plants and microbial communities regulated by environmental factors along a successional gradient, which suggests strong interdependencies between the taxa. As a next step, approaches like the one described here may be used to identify indicator species and environmental variables that inform best about the diversity and composition of ecosystems, which facilitates monitoring and conservation efforts. Additionally, climate warming demands the prediction of ecosystem-wide responses and our data presents existing patterns and offers information for future predictions.

Acknowledgement

We thank the Hohe Tauern National Park Salzburg Administration and the Rudolfshütte for organizational and logistic support, the governing authority Land Salzburg for the permit to conduct our research (permit # 20507-96/45/7-2019). Jan-Christoph Otto, Tobias Seifert, and Anna Vojtkó for help in the field. Hamed Azarbad, Lisa-Maria Ohler and Verena Zieschank provided valuable comments to improve the study. The study was supported by the START-program of the Austrian Science Fund (FWF) granted to Robert R. Junker (Y1102).

Author contributions

RRJ conceived the study. XH, MH, VRH and RRJ designed the study and collected the data. XH and RRJ analyzed the data. XH and RRJ wrote the manuscript with critical input from MH and VRH. All authors contributed to the manuscript and approved the final version.

References

Aguiar FC, Cerdeira JO, Martins MJ, Ferreira MT & Pillar V. 2013. Riparian forests of Southwest Europe: are functional trait and species composition assemblages constrained by environment? *Journal of Vegetation Science* **24**: 628-638.

Angly FE, Dennis PG, Skarshewski A, Vanwonterghem I, Hugenholtz P & Tyson GW. 2014. CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome* **2**: 11.

Bardgett RD, Richter A, Bol R, Garnett MH, Baumler R, Xu X, Lopez-Capel E, Manning DA, Hobbs PJ, Hartley IR et al. 2007. Heterotrophic microbial communities use ancient carbon following glacial retreat. *Biology Letters* **3**: 487-490.

Berg G, Rybakova D, Fischer D, Cernava T, Verges MC, Charles T, Chen X, Cocolin L, Eversole K, Corral GH et al. 2020. Microbiome definition re-visited: old concepts and new challenges. *Microbiome* **8**: 103.

Bernhardt-Römermann M, Römermann C, Nuske R, Parth A, Klotz S, Schmidt W &

458 Stadler J. 2008. On the identification of the most suitable traits for plant
459 functional trait analyses. *Oikos* **117**: 1533-1541.
460 Bever JD, Platt TG & Morton ER. 2012. Microbial population and community
461 dynamics on plant roots and their feedbacks on plant communities. *Annual*
462 *Review of Microbiology* **66**: 265-283.
463 Biau G & Scornet E. 2016. A random forest guided tour. *Test* **25**: 197-227.
464 Blanchet FG, Cazelles K & Gravel D. 2020. Co-occurrence is not evidence of
465 ecological interactions. *Ecology Letters* **23**: 1050-1063.
466 Breiman L. 2001. Random Forests. *Machine Learning* **45**: 5-32.
467 Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK,
468 Fierer N, Pena AG, Goodrich JK, Gordon JI et al. 2010. QIIME allows analysis
469 of high-throughput community sequencing data. *Nature Methods* **7**: 335-336.
470 Chang C & HilleRisLambers J. 2016. Integrating succession and community assembly
471 perspectives. *F1000Research* **5**: 1-10.
472 Chen W, Xu R, Wu Y, Chen J, Zhang Y, Hu T, Yuan X, Zhou L, Tan T & Fan J. 2018.
473 Plant diversity is coupled with beta not alpha diversity of soil fungal
474 communities following N enrichment in a semi-arid grassland. *Soil Biology and*
475 *Biochemistry* **116**: 388-398.
476 Chen Y-L, Xu T-L, Veresoglou SD, Hu H-W, Hao Z-P, Hu Y-J, Liu L, Deng Y, Rillig
477 MC & Chen B-D. 2017. Plant diversity represents the prevalent determinant of
478 soil fungal community structure across temperate grasslands in northern China.
479 *Soil Biology and Biochemistry* **110**: 12-21.
480 Cheng J, Zhao M, Cong J, Qi Q, Xiao Y, Cong W, Deng Y, Zhou J & Zhang Y. 2020.
481 Soil pH exerts stronger impacts than vegetation type and plant diversity on soil
482 bacterial community composition in subtropical broad-leaved forests. *Plant and*
483 *Soil* **450**: 273-286.
484 Darcy JL, Schmidt SK, Knelman JE, Cleveland CC, Castle SC & Nemergut DR. 2018.
485 Phosphorus, not nitrogen, limits plants and microbial primary producers
486 following glacial retreat. *Science Advances* **4**: eaaq0942.

487 Dassen S, Cortois R, Martens H, de Hollander M, Kowalchuk GA, van der Putten WH
488 & De Deyn GB. 2017. Differential responses of soil bacteria, fungi, archaea and
489 protists to plant species richness and plant functional group identity. *Molecular*
490 *Ecology* **26**: 4085-4098.
491 David AS, Thapa-Magar KB, Menges ES, Searcy CA & Afkhami ME. 2020. Do plant-
492 microbe interactions support the Stress Gradient Hypothesis? *Ecology* **101**:
493 e03081.
494 Davison J, Moora M, Semchenko M, Adenan SB, Ahmed T, Akhmetzhanova AA,
495 Alatalo JM, Al-Quraishy S, Andriyanova E, Anslan S et al. 2021. Temperature
496 and pH define the realised niche space of arbuscular mycorrhizal fungi. *New*
497 *Phytologist* **231**: 763-776.
498 Deveau A, Bonito G, Uehling J, Paoletti M, Becker M, Bindschedler S, Hacquard S,
499 Herve V, Labbe J, Lastovetsky OA et al. 2018. Bacterial-fungal interactions:
500 ecology, mechanisms and challenges. *FEMS Microbiology Reviews* **42**: 335-352.
501 Dixon P. 2003. VEGAN, a package of R functions for community ecology. *Journal of*
502 *Vegetation Science* **14**: 927-930.
503 Fischer HS, Michler B, Ziche D & Fischer A. 2019. Plants as Indicators of Soil
504 Chemical Properties. In: Wellbrock N, Bolte A, eds. *Status and Dynamics of*
505 *Forests in Germany*. Ecological Studies (Analysis and Synthesis), Springer,
506 Cham, 295-309.
507 Ghannam RB & Techtmann SM. 2021. Machine learning applications in microbial
508 ecology, human microbiome studies, and environmental monitoring.
509 *Computational and Structural Biotechnology Journal* **19**: 1092-1107.
510 Goodswen SJ, Barratt JLN, Kennedy PJ, Kaufer A, Calarco L & Ellis JT. 2021. Machine
511 learning and applications in microbiology. *FEMS Microbiology Reviews*. doi:
512 10.1093/femsre/fuab015
513 Hahl T, Moorsel SJ, Schmid MW, Zuppinger-Dingley D, Schmid B, Wagg C & Thakur
514 M. 2020. Plant responses to diversity-driven selection and associated
515 rhizosphere microbial communities. *Functional Ecology* **34**: 707-722.

516 Harrison S, Spasojevic MJ & Li D. 2020. Climate and plant community diversity in
517 space and time. *Proceedings of the National Academy of Sciences* **117**: 4464-
518 4470.
519 Hartley IP, Heinemeyer A, Evans SP & Ineson P. 2007. The effect of soil warming on
520 bulk soil vs. rhizosphere respiration. *Global Change Biology* **13**: 2654-2667.
521 Heinen R, Hannula SE, De Long JR, Huberty M, Jongen R, Kielak A, Steinauer K, Zhu
522 F & Bezemer TM. 2020. Plant community composition steers grassland
523 vegetation via soil legacy effects. *Ecology Letters* **23**: 973-982.
524 Hermans SM, Buckley HL, Case BS, Curran-Cournane F, Taylor M & Lear G. 2020.
525 Using soil bacterial communities to predict physico-chemical variables and soil
526 quality. *Microbiome* **8**: 79.
527 Hintze C, Heydel F, Hoppe C, Cunze S, König A & Tackenberg O. 2013. D3: The
528 Dispersal and Diaspore Database – Baseline data and statistics on seed dispersal.
529 *Perspectives in Plant Ecology, Evolution and Systematics* **15**: 180-192.
530 Horn S, Hempel S, Verbruggen E, Rillig MC & Caruso T. 2017. Linking the community
531 structure of arbuscular mycorrhizal fungi and plants: a story of interdependence?
532 *The ISME Journal* **11**: 1400-1411.
533 Junker RR, Hanusch M, He X, Ruiz-Hernández V, Otto J-C, Kraushaar S, Bauch K,
534 Griessenberger F, Ohler L-M & Trutschnig W. 2020. Ödenwinkel: an Alpine
535 platform for observational and experimental research on the emergence of
536 multidiversity and ecosystem complexity. *Web Ecology* **20**: 95-106.
537 Kahmen S & Poschlod P. 2004. Plant functional trait responses to grassland succession
538 over 25 years. *Journal of Vegetation Science* **15**: 21-32.
539 Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD,
540 Blomberg SP & Webb CO. 2010. Picante: R tools for integrating phylogenies
541 and ecology. *Bioinformatics* **26**: 1463-1464.
542 Knelman JE, Legg TM, O'Neill SP, Washenberger CL, González A, Cleveland CC &
543 Nemergut DR. 2012. Bacterial community structure and function change in
544 association with colonizer plants during early primary succession in a glacier

545 forefield. *Soil Biology and Biochemistry* **46**: 172-180.

546 Kruger C, Kohout P, Janouskova M, Puschel D, Frouz J & Rydlova J. 2017. Plant
547 Communities Rather than Soil Properties Structure Arbuscular Mycorrhizal
548 Fungal Communities along Primary Succession on a Mine Spoil. *Frontiers in*
549 *Microbiology* **8**: 719.

550 Laliberte E & Legendre P. 2010. A distance-based framework for measuring functional
551 diversity from multiple traits. *Ecology* **91**: 299-305.

552 Lange M, Eisenhauer N, Sierra CA, Bessler H, Engels C, Griffiths RI, Mellado-
553 Vazquez PG, Malik AA, Roy J, Scheu S et al. 2015. Plant diversity increases
554 soil microbial activity and soil carbon storage. *Nature Communications* **6**: 6707.

555 Leff JW, Bardgett RD, Wilkinson A, Jackson BG, Pritchard WJ, De Long JR, Oakley
556 S, Mason KE, Ostle NJ, Johnson D et al. 2018. Predicting the structure of soil
557 communities from plant community taxonomy, phylogeny, and traits. *The ISME*
558 *Journal* **12**: 1794-1805.

559 Lopez-Angulo J, de la Cruz M, Chacon-Labella J, Illuminati A, Matesanz S, Pescador
560 DS, Pias B, Sanchez AM & Escudero A. 2020. The role of root community
561 attributes in predicting soil fungal and bacterial community patterns. *New*
562 *Phytologist* **228**: 1070-1082.

563 Losapio G, Cerabolini BEL, Maffioletti C, Tampucci D, Gobbi M & Caccianiga M.
564 2021. The Consequences of Glacier Retreat Are Uneven Between Plant Species.
565 *Frontiers in Ecology and Evolution* **8**:

566 Miller TK, Heegaard E, Hassel K, Kapfer J & Bartha S. 2020. Environmental variables
567 driving species composition in subarctic springs in the face of climate change.
568 *Journal of Vegetation Science* **32**: e12955.

569 Miransari M. 2011. Interactions between arbuscular mycorrhizal fungi and soil bacteria.
570 *Applied Microbiology and Biotechnology* **89**: 917-930.

571 Mitchell RJ, Hester AJ, Campbell CD, Chapman SJ, Cameron CM, Hewison RL &
572 Potts JM. 2011. Explaining the variation in the soil microbial community: do
573 vegetation composition and soil chemistry explain the same or different parts of

574 the microbial variation? *Plant and Soil* **351**: 355-362.

575 Mony C, Gaudu V, Ricono C, Jambon O & Vandenkoornhuyse P. 2021. Plant
576 neighbours shape fungal assemblages associated with plant roots: A new
577 understanding of niche-partitioning in plant communities. *Functional Ecology*
578 **35**: 1768-1782.

579 Mouhamadou B, Puissant J, Personeni E, Desclos-Theveniau M, Kastl EM, Schloter M,
580 Zinger L, Roy J, Geremia RA & Lavorel S. 2013. Effects of two grass species
581 on the composition of soil fungal communities. *Biology and Fertility of Soils*
582 **49**: 1131-1139.

583 Navratilova D, Tlaskalova P, Kohout P, Drevojan P, Fajmon K, Chytry M & Baldrian P.
584 2019. Diversity of fungi and bacteria in species-rich grasslands increases with
585 plant diversity in shoots but not in roots and soil. *FEMS Microbiology Ecology*
586 **95**:

587 Ohler LM, Lechleitner M & Junker RR. 2020. Microclimatic effects on alpine plant
588 communities and flower-visitor interactions. *Scientific Reports* **10**: 1366.

589 Ortiz-Alvarez R, Fierer N, de Los Rios A, Casamayor EO & Barberan A. 2018.
590 Consistent changes in the taxonomic structure and functional attributes of
591 bacterial communities during primary succession. *The ISME Journal* **12**: 1658-
592 1667.

593 Pearse WD, Cadotte MW, Cavender-Bares J, Ives AR, Tucker CM, Walker SC &
594 Helmus MR. 2015. pez: phylogenetics for the environmental sciences.
595 *Bioinformatics* **31**: 2888-2890.

596 Pichler M, Boreux V, Klein AM, Schleuning M, Hartig F & Carvalheiro L. 2019.
597 Machine learning algorithms to infer trait-matching and predict species
598 interactions in ecological networks. *Methods in Ecology and Evolution* **11**: 281-
599 293.

600 Porazinska DL, Farrer EC, Spasojevic MJ, Bueno de Mesquita CP, Sartwell SA, Smith
601 JG, White CT, King AJ, Suding KN & Schmidt SK. 2018. Plant diversity and
602 density predict belowground diversity and function in an early successional

603 alpine ecosystem. *Ecology* **99**: 1942-1952.

604 Prober SM, Leff JW, Bates ST, Borer ET, Firn J, Harpole WS, Lind EM, Seabloom EW,
605 Adler PB, Bakker JD et al. 2015. Plant diversity predicts beta but not alpha
606 diversity of soil microbes across grasslands worldwide. *Ecology Letters* **18**: 85-
607 95.

608 Rasmussen PU, Bennett AE, Tack AJM & Singh B. 2019. The impact of elevated
609 temperature and drought on the ecology and evolution of plant–soil microbe
610 interactions. *Journal of Ecology* **108**: 337-352.

611 Raza W, Wang J, Jousset A, Friman VP, Mei X, Wang S, Wei Z & Shen Q. 2020.
612 Bacterial community richness shifts the balance between volatile organic
613 compound-mediated microbe-pathogen and microbe-plant interactions.
614 *Proceedings of the Royal Society B: Biological Sciences* **287**: 20200403.

615 Reese AT, Lulow K, David LA & Wright JP. 2018. Plant community and soil conditions
616 individually affect soil microbial community assembly in experimental
617 mesocosms. *Ecology and Evolution* **8**: 1196-1205.

618 Robroek BJM, Martí M, Svensson BH, Dumont MG, Veraart AJ & Jassey VEJ. 2021.
619 Rewiring of peatland plant–microbe networks outpaces species turnover. *Oikos*
620 **130**: 339-353.

621 Rudgers JA, Afkhami ME, Bell-Dereske L, Chung YA, Crawford KM, Kivlin SN,
622 Mann MA & Nuñez MA. 2020. Climate Disruption of Plant-Microbe
623 Interactions. *Annual Review of Ecology, Evolution, and Systematics* **51**: 561-
624 586.

625 Sanchez-Canizares C, Jorrin B, Poole PS & Tkacz A. 2017. Understanding the
626 holobiont: the interdependence of plants and their microbiome. *Current Opinion*
627 *in Microbiology* **38**: 188-196.

628 Schulz S, Brankatschk R, Dümig A, Kögel-Knabner I, Schloter M & Zeyer J. 2013. The
629 role of microorganisms at different stages of ecosystem development for soil
630 formation. *Biogeosciences* **10**: 3983-3996.

631 Shen C, Gunina A, Luo Y, Wang J, He JZ, Kuzyakov Y, Hemp A, Classen AT & Ge Y.

2020. Contrasting patterns and drivers of soil bacterial and fungal diversity across a mountain gradient. *Environmental Microbiology* **22**: 3287-3301.

Sweeney CJ, de Vries FT, van Dongen BE & Bardgett RD. 2021. Root traits explain rhizosphere fungal community composition among temperate grassland plant species. *New Phytologist* **229**: 1492-1507.

Teste FP, Kardol P, Turner BL, Wardle DA, Zemunik G, Renton M & Laliberte E. 2017. Plant-soil feedback and the maintenance of diversity in Mediterranean-climate shrublands. *Science* **355**: 173-176.

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

Vandenkoornhuyse P, Mahe S, Ineson P, Staddon P, Ostle N, Cliquet JB, Francez AJ, Fitter AH & Young JP. 2007. Active root-inhabiting microbes identified by rapid incorporation of plant-derived carbon into RNA. *Proceedings of the National Academy of Sciences* **104**: 16970-16975.

Wargo MJ & Hogan DA. 2006. Fungal--bacterial interactions: a mixed bag of mingling microbes. *Current Opinion in Microbiology* **9**: 359-364.

Wojcik R, Eichel J, Bradley JA & Benning LG. 2021. How allogenic factors affect succession in glacier forefields. *Earth-Science Reviews* **218**:

Zak DR, Holmes WE, White DC, Peacock AD & Tilman D. 2003. Plant Diversity, Soil Microbial Communities, and Ecosystem Function: Are There Any Links? *Ecology* **84**: 2042-2050.

Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, FitzJohn RG, McGlinn DJ, O'Meara BC, Moles AT, Reich PB et al. 2014. Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**: 89-92.

Figures

Fig. 1 Random forest predictions using the community table of soil bacterial communities (OTU table) to predict seven variables of plant (green) and three variables of fungi (orange) as well as seven variables of environmental factors (blue). The colored circles at the left of each plot denote the mean ± standard deviation of the accuracy of prediction using the full dataset (results of 10-fold cross validation), and the black circles denote the mean ± standard deviation of the accuracy of prediction for all the frames. Each grey to black circle on the right of each plot represents the mean accuracy of prediction of each frame and the color gradient is showing how many correlations of the 10-fold cross-validation were significant with lighter colors indicating less frequent significant predictions. A quadratic or linear regression (the model with higher adjusted r2 value) is fit for the gradient if it is significant, showing a change of the accuracy of prediction along the successional gradient.

Fig. 2 Random forest predictions using the community table of soil fungal communities (OUT table) to predict seven variables of plant (green) and three variables of bacteria (red) as well as seven variables of environmental factors (blue). The colored circles at the left of each plot denote the mean ± standard deviation of the accuracy of prediction using the full dataset (results of 10-fold cross validation), and the black circles denote the mean ± standard deviation of the accuracy of prediction for all the frames. Each grey to black circle on the right of each plot represents the mean accuracy of prediction of each frame and the color gradient is showing how many correlations of the 10-fold cross-validation were significant with lighter colors indicating less frequent significant predictions. A quadratic or linear regression (the model with higher adjusted r2 value) is fit for the gradient if it is significant, showing a change of the accuracy of prediction along the successional gradient.

Fig. 3 Random forest predictions using the community table of plant communities to predict three variables of bacteria (red) and three variables of fungi (orange) as well as seven variables of environmental factors (blue). The colored circles at the left of each plot denote the mean ± standard deviation of the accuracy of prediction using the full

dataset (results of 10-fold cross validation), and the black circles denote the mean ± standard deviation of the accuracy of prediction for all the frames. Each grey to black circle on the right of each plot represents the mean accuracy of prediction of each frame and the color gradient is showing how many correlations of the 10-fold cross-validation were significant with lighter colors indicating less frequent significant predictions. A quadratic or linear regression (the model with higher adjusted r2 value) is fit for the gradient if it is significant, showing a change of the accuracy of prediction along the successional gradient.

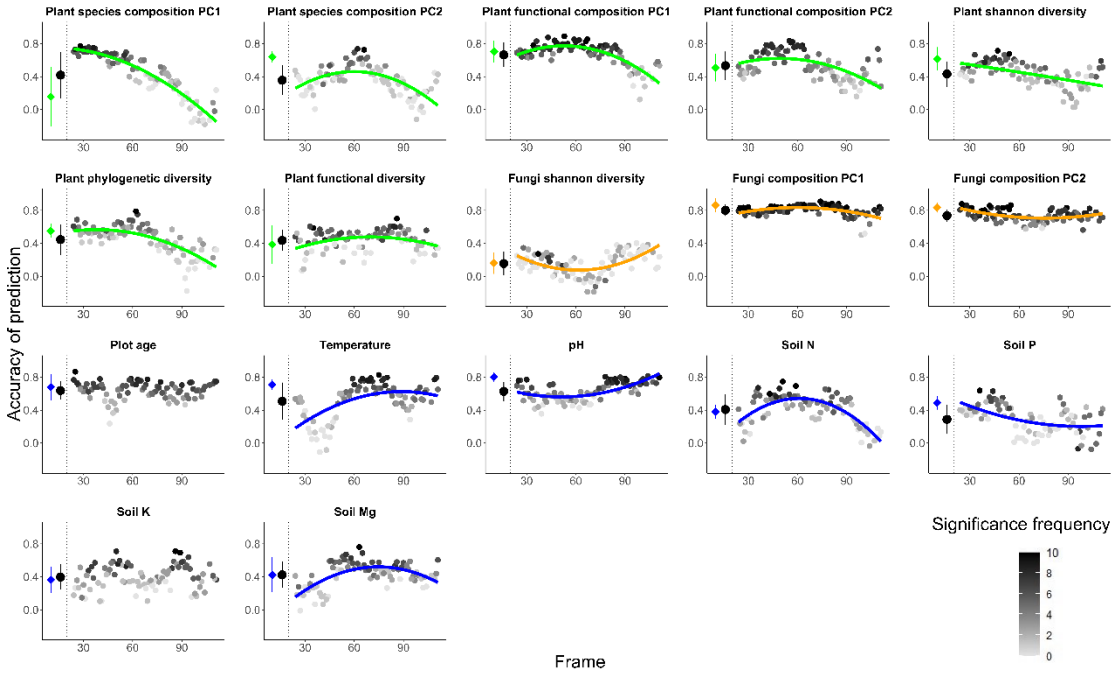
Fig. 4 Random forest predictions using all the environmental factors to predict seven variables of plant (green) and three variables of bacteria (red) as well as three variables of fungi (orange). The colored circles at the left of each plot denote the mean ± standard deviation of the accuracy of prediction using the full dataset (results of 10-fold cross validation), and the black circles denote the mean ± standard deviation of the accuracy of prediction for all the frames. Each grey to black circle on the right of each plot represents the mean accuracy of prediction of each frame and the color gradient is showing how many correlations of the 10-fold cross-validation were significant with lighter colors indicating less frequent significant predictions. A quadratic or linear regression (the model with higher adjusted r2 value) is fit for the gradient if it is significant, showing a change of the accuracy of prediction along the successional gradient.

Fig. 5 Summary of the accuracy of prediction using taxonomic groups (bacteria (a), fungi (b), plant (c)) and environmental factors (d) to predict variables from the other three groups along the successional gradient. Variables from each group are color-coded (red: bacteria, orange: fungi, green: plant, blue: environment) and ranked by accuracy of prediction.

Fig. 6 Summary of the accuracy of prediction for each variable being predicted by every single group (red: bacteria, orange: fungi, green: plant, blue: environment) as well as by the other three groups combined (grey). The label on each boxplot is the result of Tukey Test showing if there is significant difference of accuracy of prediction between

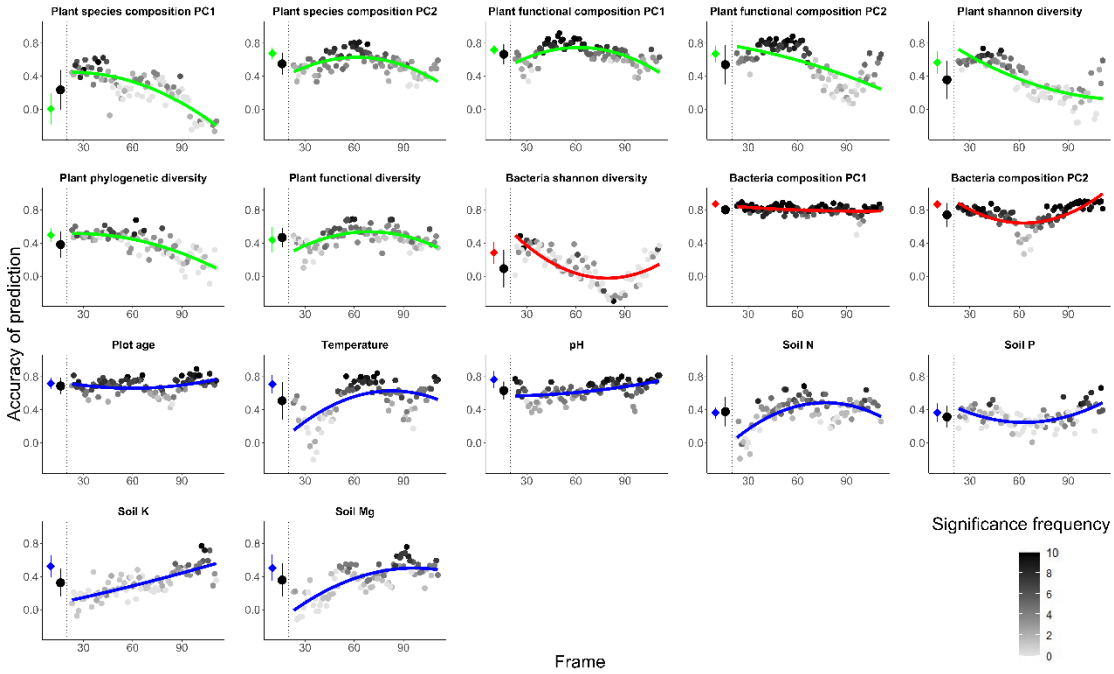
715 any pair of predicting groups.

716 **Fig. 1**



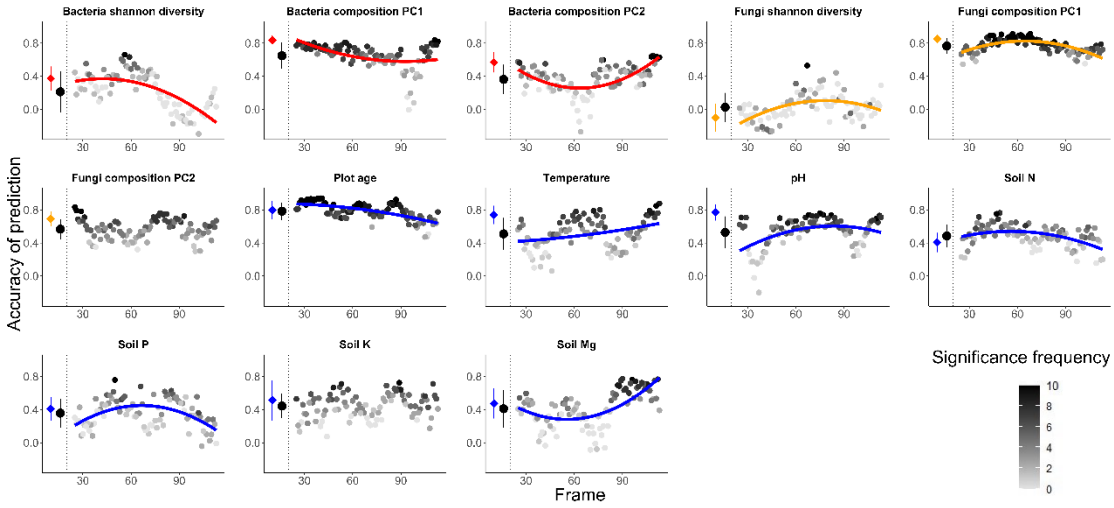
717

718 **Fig. 2**



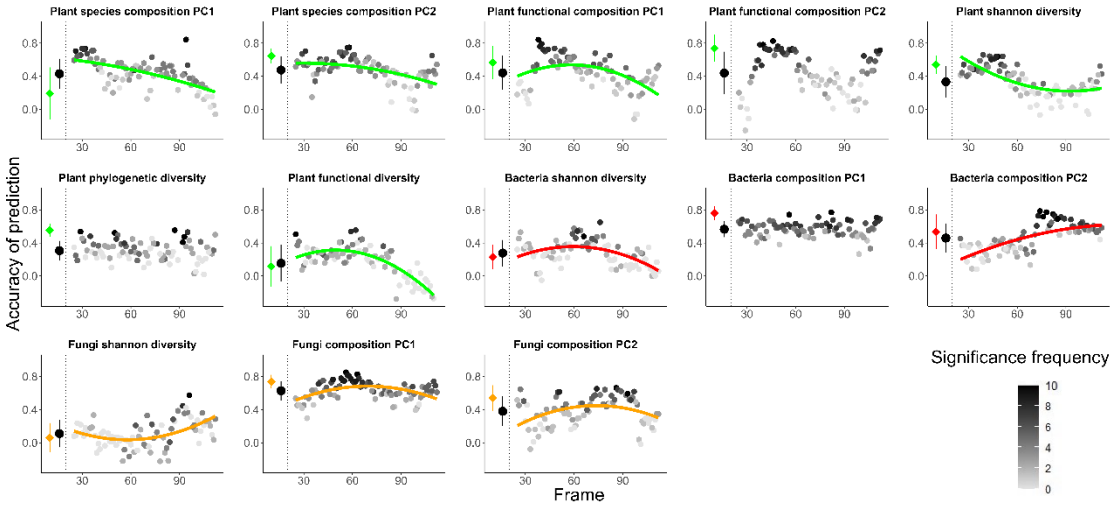
719

720 **Fig. 3**



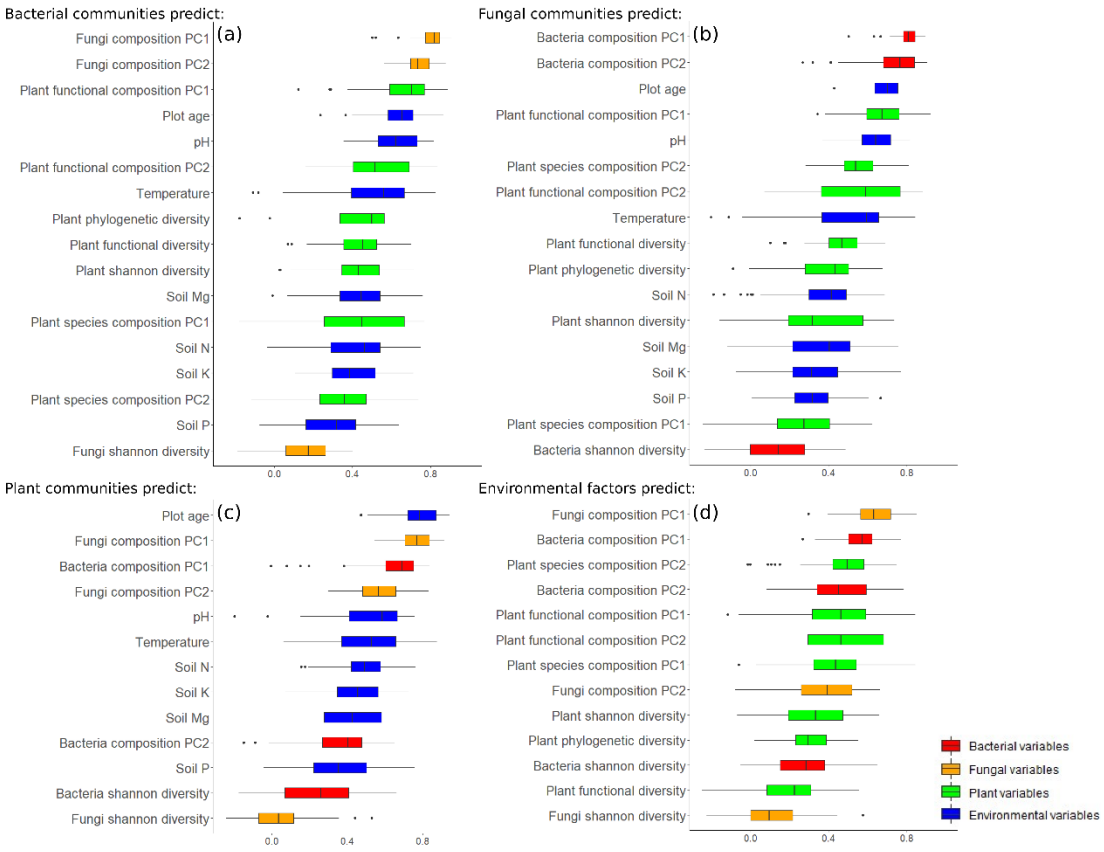
721

722 **Fig. 4**



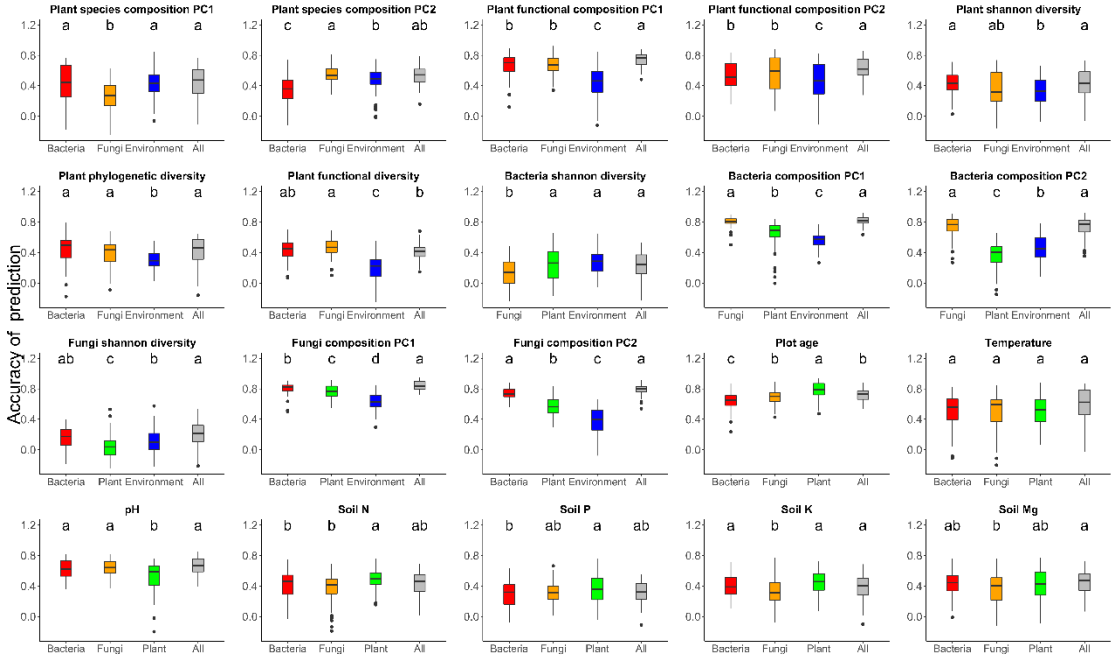
723

724 **Fig. 5**



725

Fig. 6



Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Table S1 Species composition of plant communities used as explanatory variables in predictive models.

Table S2 Community table of soil bacterial communities (OTU table) used as explanatory variables in predictive models.

Table S3 Community table of soil fungal communities (OTU table) used as explanatory variables in predictive models.

Table S4 Environmental parameters table used as explanatory variables in predictive models.

Table S5 All the 20 dependent variables used for predictive models.

Table S6 Functional composition of plant communities.

Table S7 Normalized mean value of field-measured traits (plant height, leaf area, leaf weight, SLA) for 48 species.