Soil microbial communities influencing organic phosphorus mineralization in a coastal dune chronosequence in New Zealand

Jonathan R. Gaiero, Micaela Tosi, Elizabeth Bent, Gustavo Boitt, Kamini Khosla, Benjamin L. Turner, Alan E. Richardson, Leo M. Condron, Kari E. Dunfield*

Gaiero, Jonathan; School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada
Tosi, Micaela; School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada
Bent, Elizabeth; School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada
Boitt, Gustavo; School of Agriculture and Environment, The University of Western Australia, Crawley, WA, Australia
Khosla, Kamini; School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada
Turner, Benjamin; Smithsonian Tropical Research Institute, Balboa, Ancón, Republic of Panama
Richardson, Alan; CSIRO Agriculture and Food, Canberra, ACT, Australia
Condron, Leo; Faculty of Agriculture & Life Sciences, Lincoln University, Lincoln, Canterbury, New Zealand
Dunfield, Kari; School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada

* Corresponding author. E-mail: dunfield@uoguelph.ca

Keywords

Haast chronosequence, pedogenesis, phosphorus cycling, phosphatase, phytase, phosphonatase

Abstract

The Haast chronosequence in New Zealand is a ~6500-year dune formation series, characterized by rapid podzol development, phosphorus (P) depletion, and a decline in aboveground biomass. We examined bacterial and fungal community composition within mineral soil fractions using amplicon-based high-throughput sequencing (Illumina MiSeq). We targeted bacterial non-specific acid (class A, *phoN*/*phoC*) and alkaline (*phoD*) phosphomonoesterase genes and quantified specific genes and transcripts using real-time PCR. Soil bacterial diversity was greatest after 4000 years of ecosystem development and associated with an increased richness of phylotypes and a significant decline in previously dominant taxa (Firmicutes and Proteobacteria). Soil fungal communities transitioned from predominantly Basidiomycota to Ascomycota along the chronosequence and were most diverse in 290 to 392-year-old soils, coinciding with maximum tree basal area and organic P accumulation. The Bacteria:Fungi ratio decreased amid a competitive and interconnected soil community as determined by network analysis. Overall, soil microbial communities were associated with soil changes and declining P throughout pedogenesis and ecosystem succession. We identified an increased dependence on organic P mineralization, as found by the profiled acid phosphatase genes, soil acid phosphatase activity, and function inference from predicted metagenomes (PICRUSt2).

1. Introduction

Soil chronosequences provide long-term natural experiments that offer a 'space for time substitution' and allow a unique opportunity for investigating microbial dynamics with regard to long term soil development (Wardle *et al.* 2012). Ecosystem development varies predictably along chronosequences worldwide, with early increases in plant diversity and biomass, termed

progression, followed by a decline in plant biomass linked to phosphorus (P) depletion, termed retrogression (Peltzer *et al.* 2010). The older, more nutrient depleted, soils often promote diverse forest communities (Laliberté *et al.* 2013), comprised of stress-tolerant plants, followed by further ecosystem productivity decline (Peltzer *et al.* 2010). According to Walker and Syers (1976), nitrogen (N) supports limited plant productivity on young soils, because N is largely absent from parent material and is introduced into the system by biological fixation and atmospheric deposition (Peltzer *et al.* 2010). However, P is derived primarily from parent material and thus depending on the nature of the material, it is limited, or its concentration in the soil declines continually throughout pedogenesis (Augusto *et al.* 2017). This developing P limitation impacts system productivity as soils age and provides a strong conceptual framework within which to evaluate the influence of long-term changes in nutrient availability on biological communities and resulting feedbacks (Turner *et al.* 2013).

The importance of microbially-mediated organic P cycling is widely acknowledged (Allison *et al.* 2007; Richardson and Simpson 2011; Tamburini *et al.* 2012; Turner *et al.* 2013). Previous studies of the Franz Josef and Waitutu chronosequences in New Zealand revealed significant change in soil microbial community structure and predicted contribution to organic matter turnover and P availability through mineralization (Williamson, Wardle and Yeates 2005; Allison *et al.* 2007; Turner *et al.* 2013). Along the 120,000-year-old Franz Josef chronosequence, increasing phosphatase activity and relative efficiency (activity per unit microbial biomass), over time were associated with a decline in soil P (Allison *et al.* 2007). Throughout this chronosequence, microbial biomass P became an increasingly important component of total system biomass P (Turner *et al.* 2013). Soil microbial community changes relating to P cycling are of interest due to the long timescale of P depletion and the key role of microorganisms in P

cycling and as a P sink. However, few studies have specifically examined changes in microorganisms and microbial P-cycling genes during long-term ecosystem development.

We identified for our study that the 6,500-year-old Haast sand-dune chronosequence provides an ideal template to investigate these dynamic changes in P cycling with belowground biological communities. Located in the humid southwest of New Zealand, plant communities and soil chemistry at Haast change rapidly in the early stages of ecosystem development, leading to a retrogression phase in the dunes aged around 400-500 years and older (Turner *et al.* 2012a, 2012b; Turner, Wells and Condron 2014). Previous research at Haast highlighted some of the important roles filled by microbial communities in P cycling and in connection with ecosystem development. Using a 454-pyrosequencing approach, Jangid *et al.* (2013a) showed long-term changes in bacterial phyla abundance and community structure, and using oxygen isotopes, Roberts *et al.* (2015) showed the soil microbial biomass is continuously and extensively cycling labile and bioavailable forms of P. As such, we hypothesize an increased promotion of microbial genes and activity relating to organic P mineralization as soils aged. The genetic potential for P mineralization was targeted using a combination of high-throughput sequencing (HTS) and real-time PCR for five newly developed gene targets (Ragot, Kertesz and Bünemann 2015; Bergkemper *et al.* 2016a; Fraser *et al.* 2017; Gaiero *et al.* 2018), and by using predictive functional analyses (PICRUSt2) based on 16S rRNA amplicons. We also expand our understanding of the changes occurring within the soil bacterial and fungal communities using an HTS approach. Prior to this work, fungal communities had not been characterized at Haast. Regarding overall microbial community changes, we also hypothesize there to be an increased predominance of fungi as the competition for nutrients progresses with soil age. In order to alleviate community shifts related to the increasing depth of organic horizon over time, we

sampled the upper mineral fraction only. The mineral soil fraction previously showed strong correlations to changes occurring across the chronosequence, including the tree community composition and dune age (Turner *et al.* 2012b). Sampling proceeded until 3900 years before present (yBP), dune 12, whereby soil developmental changes and total P loss were in plateau (Turner *et al.* 2012a).

2. Materials and Methods

2.1 The Haast chronosequence

The Haast dune system is a foredune progradation sequence located a few kilometers northeast of the town of Haast, on the west coast of the South Island of New Zealand (approximate center of the sequence 43°49′20″ S, 169°4′30″ E) (Figure 1A). It comprises a total of seventeen dune ridges that occur as generally continuous features across the length of the system, ranging in age from 180 years to ~6,500 years old (Wells and Goff 2007; Turner *et al.* 2012a). Dune building episodes are thought to have resulted from earthquakes along the Alpine Fault and/or prolonged periods of high rainfall, causing widespread land disturbance including landslides, rock falls, floodplain aggradation, and tree falls (Wells and Goff 2007). The Haast system approximates an ideal chronosequence, because four of the five soil forming factors (parent material, climate, vegetation, and topography) are well constrained (Turner *et al.* 2012a).

2.2 Summary of ecosystem development and soil properties

The soils across the Haast chronosequence develop rapidly to Spodosols (podzols) (Turner *et al.* 2012a) (Figure 1). Based on soil development, dunes are sub-classified as young (< 800 yBP), intermediate (800-3900 yBP), and old (3900-6500 yBP). Pedogenesis at Haast involves an increase in the organic horizon depth that occurs at up to 28 cm on the older dunes, and increasingly acidified sandy and loamy sand mineral soils (pH 4.5 to 3.5) (Table 1). The

chronosequence is also characterized by a rapid decline of total P in the first few hundred years of pedogenesis, and with an organic fraction that comprises up to 85% of total P (Figure 1C). Declining P, with concomitant increasing total C and N, results in a significant increase in C:P and N:P ratios, indicative of strengthening P limitation as the soils age. As a result, historical measurements indicated increased activity of phosphomonoesterases and phosphodiesterases along the chronosequence (Figure 1D).

Forest structure on the youngest dunes is dominated by conifers in the Podocarpaceae family (e.g. *Dacrydium cupressinum*), peaking in basal area by dune 4 (517 yBP). On the older dunes, forest structure transitions to a mixed conifer–woody angiosperm community (Turner *et al.* 2012b) (Figure 1B). Ferns (e.g. *Dicksonia squarrosa* and *Cyathea smithii*) typically occur on younger dunes and are not found on dunes older than 2000 yBP.

2.3 Soil sampling and physicochemical analysis

Six dunes (1, 2, 3, 4, 10, and 12) were sampled ranging in age from 181 yBP to approximately 3903 yBP (Table 1). Along the crest of each dune, we established six replicate plots (5×10 m) separated by ~50 m and, in each of them, we collected six soil cores (0-10 cm) from the upper mineral layer (n = 36). Hence, the removed organic soil ranged from ~5 cm in depth in the young dunes to ~25 cm in the older dunes (Turner *et al.* 2012a). Soil samples of ~2 g were placed in LifeGuard™ Soil Preservation Solution (Qiagen, Hilden, Germany) and kept cool on ice to preserve microbial community integrity during transportation. Soil samples were also brought to the laboratory, sieved (Ø = 2 mm), and stored at 4°C until further processing within 24 hours. Potential acid phosphatase activity (ACP) was measured by the release of ρ-nitrophenol (ρ-NP) after incubation of soil with ρ-nitrophenyl phosphate (ρNPP) (Tabatabai and Bremner 1969). Phosphatase activity is presented in μmol ρ-NP g^{-1} h^{-1}. Bioavailable inorganic P or Olsen

P (mg P kg^{-1}) was determined by extraction with 0.5 M $NaHCO_3$ for 30 minutes (Olsen 1954). The remaining soil samples were dried at 65 °C to constant weight for further biogeochemical characterization. Soil pH was determined in H_2O (1:2.5 v/v). Total carbon (C) and total N were measured using an elemental analyzer (LECO 2000 CNS Analyser) in triplicate.

2.4. Nucleic acid extractions

Soil samples were briefly centrifuged to decant the Lifeguard™ (Qiagen) solution. RNA and DNA were co-extracted using the PowerSoil® Total RNA Isolation Kit (Qiagen), followed by the PowerSoil® DNA Elution Accessory Kit (Qiagen), following the manufacturer's instructions. Strong PCR inhibition was identified in the eluted DNA as a result of co-extracted substances, so a secondary cleanup step was performed using the reagents from the PowerSoil® Total DNA Isolation Kit (Qiagen) and stored at -20 °C until analysis. Complementary DNA (cDNA) suitable for downstream analyses was generated from RNA samples (5 µL, in triplicate) following DNase treatment and reverse transcription (RT) using Maxima™ H Minus cDNA Synthesis Kit with dsDNase (Thermo Scientific™). No-RT controls and no-template controls were carried out to confirm the effectivity of the DNase treatment and no contamination, respectively.

2.5 Quantitative PCR

Total gene copies and total transcripts (cDNA) for bacteria (16S rRNA), fungi (18S rRNA), and bacterial genes relating to organic P cycling were quantified when possible. Bacterial non-specific acid phosphatase genes (class A: *phoC/phoN*, and class C: *olpA*/*lppC*) and transcripts were quantified as previously described (Fraser *et al.* 2017; Gaiero *et al.* 2018, 2020). The BPP (beta-propeller phytase, a type of *myo*-inositol hexakisphosphate phosphohydrolase, EC 3.1.3.8) (Huang *et al.* 2009) and phosphonate-mediating genes *phnX* (phosphonoacetaldehyde hydrolase;

EC 3.11.1.1) (Bergkemper *et al.* 2016a), and *phnA* (phosphonoacetate hydrolase, EC 3.11.1.2) (Gilbert *et al.* 2009) were adapted for use in qPCR (Supplemental Table S1).

The qPCRs were run on 96 well plates using the CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories Inc., Hercules, CA, USA). All qPCRs were run in 20 µl reactions comprising 10 µl of 2X SsoFast™ EvaGreen® Supermix (Bio-Rad), 2 µl (0.4 µM) of each primer (concentration was gene dependent; see Supplemental Table S1), 4 µl of diluted metagenomic DNA or cDNA, and nuclease free sterile water. For fungal 18S rRNA genes, iQ™ SYBR® Green Supermix (Bio-Rad) was used instead because it showed improved amplification. Optimized qPCR protocols can be found in the Supplementary materials (Table S1). Gene and transcript copies were calculated using a standard curve based on serial 10-fold dilutions of a GeneArt plasmid vector (ThermoFisher). The GeneArt vector sequence included non-degenerate primer binding sites for *phnX*, *phnA*, *phoD*, and three bacterial non-specific acid phosphatase genes (class A, B, C) (Fraser *et al.* 2017; Gaiero *et al.* 2018). No template control measurements were always 4 or more cycles higher than sample reads. The efficiency and R^2 of each reaction are presented in Supplementary materials (Table S1). Mean cDNA copy numbers per gram dry soil with standard deviation are shown when possible, as cDNA quantitation was often below detection or non-specific, where samples were removed from analysis. Positive samples that were below the limit of detection were conservatively considered zero copies.

2.6 High-throughput sequencing and bioinformatics pipeline

2.6.1. Bacterial and fungal communities

The prokaryotic 16S rRNA V4 region and fungal ITS1 were targeted for HTS using the MiSeq platform (250bp paired-end) (Genome Quebec, Montréal, Québec). The primers were chosen to

match the work by the Earth Microbiome Project (Supplemental Table S2). Raw fastq files were processed using QIIME 2 (2018.6) (Bolyen *et al.* 2019). The plugin DADA2 was used to denoise both the 16S rRNA and ITS reads and to generate amplicon sequence variants (ASVs, also referred to as 99.9% OTUs) (Callahan *et al.* 2016) (via q2-dada2). For 16S rRNA, a total of 1,402,635 (average of 39k per sample) quality-filtered and trimmed reads were generated, and 1,159,728 (average of 32k per sample) for ITS (Supplemental Table S3). Fungal sequences were first trimmed using q2-ITSxpress (v1.7.0) to avoid artifacts caused by length polymorphism (Rivers et al., 2018), which results in improved taxonomic classification. 16S rRNA reads were filtered to remove mitochondria and chloroplast, and archaeal 16S rRNA reads were analyzed separately. Bacterial and Fungal ASVs were aligned with mafft (Katoh 2002) (via q2-alignment) and used to construct a phylogeny with fasttree2 (Price, Dehal and Arkin 2010) (via q2-phylogeny). Taxonomy was assigned to ASVs using the q2-feature-classifier (Bokulich *et al.* 2018) classify-sklearn naïve Bayes taxonomy classifier against the prokaryotic reference sequences from SILVA (v 132) with 99% OTUs (Quast *et al.* 2013), and for fungi using the dynamic classifier for UNITE (v.8.0) (Nilsson *et al.* 2019) and .

2.6.2. P-cycling genes

To limit sequencing costs, two replicate DNA extractions were pooled, leaving three pooled samples per dune (n = 3; total samples = 18). PCR amplicons of the bacterial non-specific acid

(class A, *phoN/phoC*) and alkaline phosphatase (*phoD*) genes were generated, and sequenced on the MiSeq platform (Genomics Facility at the University of Guelph) as previously outlined (Gaiero *et al.* 2020) (Supplemental Table S2). Sequence reads for class A were clustered at an 88% sequence similarity (Gaiero *et al.* 2020), and at 75% for *phoD* (Tan *et al.* 2013). There was a total of 3,998,784 reads resulting in 356 operational taxonomic units (OTUs) for class A, and 94,758 reads resulting in 276 OTUs for *phoD*. For the non-specific acid (class A, *phoN/phoC*) and alkaline (*phoD*) phosphatase genes, specificity and taxonomic assignment of OTUs was performed using BLASTx (Altschul 1997), with an E-value cutoff of $1e^{-10}$ to the closest matching bacterial order (Gaiero *et al.* 2020). Phylogenetic trees were generated online using the pipeline available at http://www.phylogeny.fr/ (Dereeper *et al.* 2008). Briefly, sequenced amplicons were aligned using MUSCLE (Edgar 2004) and a neighbor joining tree was generated. The resultant Newick tree file was imported to the Interactive tree of life v3 (iTOL) for visualizations (Letunic *et al.* 2016).

2.7 Functional predictions using PICRUSt2 and FUNGuild

Changes in soil bacterial functions related to organic P mineralization were estimated using PICRUSt2, which integrates existing open-source tools to predict genomes of environmentally sampled 16S rRNA gene sequences (Douglas *et al.* 2020). PICRUSt2 is a greatly improved tool as compared to PICRUSt1, with optimized genome prediction, and more stringent pathway inference (Douglas *et al.* 2020). The PICRUSt2 pipeline was run in python using the rarefied

bacterial 16S rRNA feature table and representative sequences, and using the default NSTI cutoff of 2.0. PICRUSt2 uses EPA-ng (Barbera *et al.* 2019) and 'gappa' (Czech, Barbera and Stamatakis 2020) for phylogenetic placement of reads, 'castor' (Louca and Doebeli 2018) for hidden state predictions, and MinPath biological pathway reconstructions (Ye and Doak 2009). MetaCyc pathways (Caspi *et al.* 2018), Kegg orthologs (KOs) and enzymes involved in organic P mineralization were selected for further statistical analyses, with representative features visualized. For example, KOs were chosen that were reflective of the trend found for related KOs, such as the Phosphonate Transporter Subunits (*phnE*, *phnD*), Phosphate-Specific Transport System Subunits (*pstB*, *pstC*, *pstS*), Glycerol 3-Phosphate Transporter Subunits (*ugpB*, *ugpC*, *ugpE*), and C-P Lyase Subunits (*phnG, phnH, phnI, phnJ, phnL, phnM*). Further, PICRUSt2 also provides the ASV contribution to each predicted function, allowing taxonomy-informed analyses to be conducted.

Additionally, FUNGuild v1.1 (Nguyen *et al.* 2016) was used to understand potential changes in ecosystem function by analyzing fungal sequences by trophic guild. Classifications were used if they were considered "highly probable" or "probable", discarding "possible" classifications for greater confidence in the results. The confidence of classifications is based on primary research literature (Nguyen *et al.* 2016). As a result, we utilized 26-36% of the fungal ITS sequences in dunes 1-4 (181-517 yBP) on average, 7% in dune 10 (3000 yBP), and 18% in dune 12 (3900 yBP).

2.8 Data analysis

Rarefaction curves were generated for all sequencing libraries, identifying adequate sampling depth for bacteria and fungi, and highlighting the sensitivity of functional gene analyses to sampling depth (Supplemental Figure S1).

Alpha and beta diversity analyses of bacterial and fungal communities were carried out on rarefied ASV tables (21,459 reads and 17,302 reads for 16S rRNA and ITS, respectively). Archaeal ASVs were found only in the oldest dunes (3000 and 3900 yBP). However, utilizing Archaeal specific qPCR primers (Uksa *et al.* 2015), we found that Archaea were quite abundant throughout the chronosequence, highlighting as others have, that the universal primers are not currently able to target all prokaryotes equally (Bahram *et al.* 2019). Therefore, the Archaeal ASVs will not be included in the remainder of analyses. Alpha diversity indices such as Faith's Phylogenetic Diversity (Faith's PD) were estimated on QIIME 2. Beta diversity and community composition analyses were carried out using Bray-Curtis dissimilarities. To visualize changes in community composition between dune stages, we generated two-dimensional non-metric multidimensional scaling (NMDS) ordination plots using Canoco 5 software v5.12 (Smilauer and Lepš 2014). To test if these differences were statistically significant, PERMANOVA tests (999 permutations), global and pairwise between dune stages, were performed in the program PAST (Hammer, Harper and Ryan 2001). Canonical correspondence analysis (CCA) was used to understand how well the environmental variables (Table 1) explained the biological variation

observed in the microbial communities. Environmental variables were included using a forward-selection approach and tested with 999 permutations, using the program Canoco 5 v5.12. Alkaline phosphatase gene (*phoD*) sequence reads were not rarefied due to low read depth, and thus acid phosphatase (class A, *phoN*/*phoC*) reads were also not rarefied for consistency. Beta diversity and community composition analyses were carried out similarly to bacteria and fungi. Further, a presence/absence analysis of class A and *phoD* OTUs between the youngest and oldest dunes was carried out to determine major changes in diversity along the chronosequence.

Changes in the relative abundance of taxonomic groups and PICRUSt2-predicted Enzymes (EC) and Kegg Orthologs (KOs), expressed as centered log-ratio (clr), were analyzed using the package 'ALDEx2' (Fernandes *et al.* 2013) in R v. 3.6.3 (R Core Team 2020). Using the function *aldex*, changes between young and old dunes were evaluated using Wilcoxon Rank Sum test with Benjamini-Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg 1995). These differential abundance tests were carried out on raw abundance data tables. Predicted ECs were also correlated to changes in taxonomic abundance (16S rRNA ASVs). Pearson correlations and visualizations were made using the package 'corrplot' in R (Taiyun Wei *et al.* 2017).

Changes in quantified gene copies and environmental parameters between dunes were analyzed using one-way analysis of variance and Tukey's test for pairwise comparisons using the R package 'multcomp' (Hothorn, Bretz and Westfall 2008). To meet assumptions of normality,

data were either log_{10} transformed, asin transformed, or left untransformed. Alpha diversity, FUNGuild trophic guilds, and cDNA gene ratios were tested using the Kruskal-Wallis non-parametric test followed by pairwise comparisons with Benjamini-Hochberg FDR correction. All sequencing data were deposited in the Sequence Read Archive (SRA) of the National Centre for Biotechnology Information (NCBI) under BioProject ID PRJNA657880.

2.9 – Co-occurrence networks

For network analysis, bacterial and fungal ASV tables were combined and filtered to include only those ASVs >2 average reads per sample (Friedman and Alm 2012). In order to improve the strength of the analysis, dunes representing the different segments of the transect were pooled as follows: young (181-290 yBP), intermediate (392-517 yBP), and old (3000-3900 yBP). The 500 most abundant ASVs were then analyzed for co-occurrence using SparCC (Friedman and Alm 2012), designed for compositional data, on the Galaxy/IDENA pipeline (http://mem.rcees.ac.cn:8081/) (Afgan *et al.* 2018; Feng *et al.* 2019). Networks included positive and negative correlations, using a cutoff of $-0.6 > R > 0.6$ ($P < 0.01$). Network topological properties and visualization was calculated using Gephi 0.9.2 (https://gephi.org/) (Bastian, Heymann and Jacomy 2009), with layouts based on the Fruchterman-Reingold algorithm and with non-overlapping nodes. Putative 'hub' or 'keystone' species (ASVs) were estimated from maximal betweenness centrality scores, relating to their importance in maintaining network

connectivity (Vick-Majors, Priscu and A Amaral-Zettler 2014; Banerjee *et al.* 2016; Reji *et al.* 2019).

3. Results

3.1 Microbial succession through the chronosequence

3.1.1 Abundance and diversity

Total bacterial abundance was relatively constant until sometime between 500 and 3000 yBP, when it declined two orders of magnitude as compared to 181 yBP ($P < 0.05$) (Figure 2A). Fungal abundance tended to increase with dune age but was not significantly different (Figure 2B). As a result, the soil microbial community transitioned to being less bacterial dominant, with a significant decline in the Bacteria:Fungi ratio with dune age (Table 1).

Bacterial and fungal communities at the ASV level changed significantly across the dune chronosequence ($P < 0.001$) (Figure 3A, B). A transitionary period was observed around 392-517 yBP, with the largest changes occurring between young dunes (181-290 yBP) and older dunes (3000-3900 yBP) (Supplemental Tables S4 and S5). Measures of bacterial alpha diversity were significantly influenced by dune age and were generally higher in the older dunes (3000-3900 yBP) ($P < 0.05$), but were relatively unchanged before 517 yBP (dunes 1-4) (Figure 2C, Supplemental Figure S2). Fungal alpha diversity, ASV richness and Faith's PD, was influenced by dune age but did not significantly change after 3900 years of development (Figure 2D, Supplemental Figure S2). Constrained analyses revealed that four soil variables (C:N, total C, Olsen P, and pH) explained a significant but small component of the variation observed in the microbial communities (Table 1, Supplemental Figure S9). Most of the variation was not explained by any combination of soil variables. The C:N ratio, total C, and pH, accounted for

7.4%, 4.1%, and 3.9% of the variation in bacterial communities, respectively (Supplemental Figure S8). For fungi, the C:N ratio was the most significant, followed by pH and Olsen P, explaining 4.6%, 3.8%, and 3.8% of the variation, respectively. This is consistent with increasing C:N ratio amid a declining pH gradient across the chronosequence (Table 1). By contrast, Olsen P and ACP, while variable across the dunes, did not differ significantly across the chronosequence.

3.1.2 Taxonomic shifts

Bacterial succession can be summarized by the predominance of Firmicutes and Proteobacteria in the youngest dunes, transitioning between 500 and 3000 yBP to a broader mixture of phyla on the older dunes, including Acidobacteria, Actinobacteria, Planctomycetes, and Verrucomicrobia (Figure 2E; Supplemental Figure S3). Overall, the large decline in the dominant phyla, particularly Firmicutes explained much of the decline in total bacterial abundance (Figure 2A), which was not compensated by substantial increases in relative abundance of other phyla (Figure 2E).

The succession of fungal communities involved a significant shift from Basidiomycota to Ascomycota-dominance over time (Figure 2F). At the order level, this succession was seen primarily as a transition from a Trichosporonales-dominated to a Hypocreales-dominated community. The transition period between 290-517 yBP saw the highest level of fungal diversity, followed by declining diversity into the oldest dunes. This was due to a decline in abundance of rarer phyla, including Mortierellomycota and Rozellomycota, along with increasing abundance of Ascomycota. In terms of functionality, the abundance of potentially pathotrophic fungi decreased over time, while symbiotrophic and saprotrophic fungi were

relatively stable over time, except for a generalized increase in saprotrophs between 290-517 yBP, peaking around 392 yBP (Supplemental Figure S7).

3.1.3 Community interactions visualized by co-occurrence networks

Co-occurrence networks were generated to identify structural community changes occurring over time (Figure 4). Network topological properties suggest that, through ecosystem development and pedogenesis, microbial communities showed reduced modularity, with more closely connected networks in the oldest dunes (lower average network distance, i.e reduced path length) (Table 2). The number of edges increased in the latter stages of the chronosequence with a shift from fungal-fungal (F-F) to bacterial-bacterial (B-B) and bacterial-fungal (B-F) connections. The number of fungal nodes decreased from 55-60% in the young and intermediate networks to 27% in the oldest, while the percentage of negative connections increased in the older dunes, from 14.6% to 36.8%, mostly due to a large increase in B-F connections and as the percentage of B-B negative connections decreased. Putative 'hub' (keystone) species estimated from each network were mainly bacterial ASVs and comprised a number of uncultured or unclassified genera (Supplemental Table S9). Overall, the declining betweenness centrality of the top keystone ASVs, and increasing average betweenness centrality per node (526 in young, 545 in intermediate, to 809 in old dunes) indicates a diminishing reliance on a few specific nodes controlling the flow of connections in the network.

3.2 Increasing relative abundance of P-cycling genes amid declining bacterial dominance

3.2.1 Abundance and gene expression of P cycling genes

Relative to the bacterial 16S rRNA gene (Figure 2A), class A acid phosphatase and *phoD* alkaline phosphatase increased in abundance 3.79 and 2.24-fold, respectively, from 181 to 3900 yBP ($P < 0.01$) (Figure 5, Supplemental Figure S4). Class C acid phosphatase also increased,

particularly up to 517 yBP ($P < 0.05$), but was not consistently different throughout the chronosequence due to high variability in the older dunes. The relative abundance of BPP and the phosphonate cycling genes (*phnX* and *phnA*) varied through the chronosequence, but did not significantly change over time, either due to high variability or increasing 16S rRNA gene copies (Figure 5, Supplemental Figure S4).

Gene transcripts were quantified for all targets except BPP (Supplemental Figure S), confirming that these genes were being expressed across the chronosequence. In absolute abundance, there were few to no changes in transcript abundances (Supplemental Figure S). In contrast, when compared to 16S rRNA cDNA, the relative abundance of class A and class C acid phosphatase appeared to increase 4 to 5 log, from 181 to 3900 yBP (data not shown). However, since many samples had low abundance or targets below detection, no statistical differences were found (data not shown).

3.2.2 Functional inferences (PICRUSt2)

PICRUSt2 predicted P-cycling genes (Kegg orthologs) and enzymes (grouped by EC) also increased in abundance over time (Figure 6). This included genes and enzymes relating to phosphate transporters, regulators, and degradation genes, which commonly showed significantly greater abundance in the older dunes (Figure 6). Predicted acid and alkaline (*phoD*) phosphatases both showed similar positive trends as found by qPCR, increasing over time, but particularly for acid phosphatase (Figure 5). Predicted genes and enzymes relating to phosphonate transport and metabolism were generally unchanged across the chronosequence ($P > 0.05$) (Figure 6), which approximates our quantitative findings for *phnX* and *phnA* (Figure 5). Further, the *phoN* acid phosphatase gene (K09474) was positively correlated with the class A gene copies as measured by qPCR (Pearson, $R^2 = 0.37$, $P < 0.0001$).

The proportions of taxonomic contribution were also estimated from the EC PICRUSt2 data (Supplemental Figure S5). Commonly, an increased contribution from Proteobacteria (orders Elsterales, Rhizobiales, Caulobacterales) and Acidobacteria (Acidobacterales) were found (Supplemental Figure S5), despite declining Proteobacterial abundance overall (Figure 2E).

3.2.3 Phosphatase-harbouring bacterial community dynamics

The acid and alkaline phosphatase gene-harbouring communities changed significantly across the chronosequence ($P < 0.001$) (Figure 7A, B); Supplemental Tables S6 and S7). A presence/absence analysis revealed a greater richness of both acid (class A) and alkaline (*phoD*) phosphatase gene OTUs in the younger dunes (181-290 yBP) when compared to the oldest dunes (3000-3900 yBP) (Figure 7C, D), highlighting a loss of unique and rare 'species'.

At the order level for class A acid phosphatase, greater abundance of Sphingomonadales was observed in the early-intermediate dunes (290-517 yBP) (Supplemental Figure S6). The older dunes revealed increasing abundance of Desulfobacterales and Syntrophobacterales (Supplemental Figure S6), orders more commonly found in anoxic environments. In several dunes, there remained a significant number of unclassified OTUs that did not meet the threshold for estimating taxonomy. Compositional succession of the *phoD* community was observable by the transition from a Pseudomonadales-dominated community to one that is Rhizobiales-dominated (Supplemental Figure S6), which may reflect changes occurring in the total bacterial community. Yet, notably, no Rhizobiales were detected within class A communities in the older dunes (Supplemental Figure S6).

4. Discussion

4.1 Soil microbial community patterns during long-term ecosystem development

Previous research at the Haast chronosequence, and other long-term chronosequences, has helped to build a framework for understanding the microbial communities as they relate to and participate in ecosystem development (Williamson, Wardle and Yeates 2005; Allison *et al.* 2007; Jangid *et al.* 2013b; Williams *et al.* 2013). At Haast, we can adapt Grime's competitor-stress-tolerator-ruderal (CSR) framework as recently outlined by Fierer (2017). For example, the periodic episodes of dune building (high disturbance), and subsequent ecosystem growth (lower stress and disturbance) and then decline (high stress), are reflected by the soil microbial communities as outlined below.

4.1.1 Early ecosystem development: Young – Intermediate Dunes (181 – 517 yBP)

Bacterial communities here are dominated by copiotrophic or ruderal taxa such as Pseudomonadales (Proteobacteria), with high catabolic activity, and spore formers such as Bacillales (Firmicutes) (Figure 2E). The transitionary period around 392 and 517 yBP is marked by compositional changes (Figure 3A, Supplemental Table S4), declining modularity and increasing connectivity (e.g. increased edges, reduced path length), as well as changes in the Bacteria:Fungi ratio (Figure 4, Table 1). The transition reveals an increasingly connected community amid an increasingly competitive environment and niche segregation (Figure 4, Table 1). Coincident with the transitionary period, we find nutrient decline, peak conversion of P into the organic pool (Eger, Almond and Condron 2011; Turner *et al.* 2012a) and increased total basal area due to increasing conifer growth (Turner *et al.* 2012b). The impact of these aboveground changes occurs at a time of highest fungal diversity (Figure 2D) and putative fungal saprotrophs (Supplemental Figure S7). Changes in fungal saprotrophs indicate important changes

to nutrient cycling and redistributions (Crowther, Boddy and Hefin Jones 2012). What remains to be determined is the level of interplay between the changes observed in the fungal communities, functionally and phylogenetically, and those changes found in the plant community, i.e. the cause and effect.

4.1.2 Late ecosystem development: Intermediate-Old Dunes (~3000 to 3900 yBP)

Microbial community succession through to 3900 yBP revealed a declining Bacteria:Fungi ratio and increasing abundance of more stress-tolerant and oligotrophic phyla, including Acidobacteria and Planctomycetes (Figure 2E). Similarly trends are found at other chronosequences under similar vegetation and climate (Allison *et al.* 2007; Jangid *et al.* 2013b; Williams *et al.* 2013). During this period, edaphic factors such as lowering pH and widening C:N remain important drivers of soil microbial community structure, but did not explain the majority of the community changes (Supplemental Figure S8). The adapted Grime's CSR framework would predict that bacterial communities in the older dunes, under increased stress, will contain slower growing heterotrophic species with an increased abundance of extracellular polymeric substances (Fierer 2017). Functional predictions using PICRUSt2 supported this hypothesis, revealing an increased abundance of pathways related to amino acid and carbohydrate biosynthesis in the oldest dune (data not shown). This may relate directly to Acidobacteria, which aim to increase their longevity amid soil stress (Kielak *et al.* 2016, 2017).

A significant increase in soil bacterial alpha diversity in the older dunes (Figure 2C) was related to declining predominance of Firmicutes and Proteobacteria and increasing richness of ASVs. Previously, two of six measures of alpha diversity (Simpson's reciprocal index and Shannon's Index) were found to decline over the 6500-year span of the Haast chronosequence (Jangid *et al.* 2013b). Our results may relate specifically to the upper mineral soil fraction or be a consequence

of the greater sampling depth of the sequencing technology used in our study (Supplemental Figure S1). Increases in soil microbial diversity could be explained both by increases in plant diversity, which was observed in this and other chronosequences (Turner *et al.* 2012b; Laliberté *et al.* 2013), as well as resource partitioning (Zhou *et al.* 2002; Kramer *et al.* 2016). Considering also higher organic matter found in the albic (E, eluvial) horizon in the older dunes (Turner *et al.* 2012a), this may have positive impact on community diversity as well.

It is widely considered that fungal-dominated food webs prevail in nutrient limiting conditions (Williamson, Wardle and Yeates 2005), and may explain the observed decrease in the Bacteria:Fungi ratio. The Bacteria:Fungi ratio is influenced by P availability (Allison *et al.* 2007), and changes in the quality and abundance of organic matter inputs from plants (Williamson, Wardle and Yeates 2005). Of note, microbial networks in older dunes (3000-3900 yBP) still relied on bacterial phylotypes as 'keystone' species (Supplemental Table S8), although in an increasingly competitive environment with fungi (Figure 4, Table 2).

At Haast, the presence of an iron pan in the older dunes (Figure 1B) could result in waterlogging and anoxic conditions following heavy rainfall, but this did not affect the plant community along the sequence (Turner *et al.* 2012b). Increasingly anoxic conditions may also be found due to depth, as the top organic horizon was ~20-25 cm in the older dunes, but only 1-6 cm in the younger dunes (Table 1). Oligotrophs found in these older dunes such as Planctomycetes and Verrucomicrobia are more suited to the anoxic conditions and reduced C availability found at depth (Bergmann *et al.* 2011), but may be playing an important role in N cycling in the older dunes via annamox (Fuerst and Sagulenko 2011). An increasing abundance of plant-associated N-fixing heterotrophs, within Rhizobiales, were also found in the older dunes (Supplemental Figure S3), which may suggest an increased functional contribution to N supply in these soils as

has been suggested at Franz Josef (Turner *et al.* 2013). In a study comparing P-rich versus P-depleted forest soils, Rhizobiales were quite abundant in the P-rich soils (Bergkemper *et al.* 2016b). This order harboured a substantial number of P-cycling genes (Bergkemper *et al.* 2016b), such as the acid and alkaline phosphomonoesterases, which we also observed at Haast for the class A and *phoD* genes (Supplemental Figures S3, S6). The authors suggested Rhizobiales to be predominantly contributing to the turnover of soil P not N fixation, which may also be true at Haast.

4.2 Microbially-driven P cycling

Coincident with the increasing proportion of phosphomonoesters as the soils aged (Turner, Wells and Condron 2014), is an increase in importance of phosphomonoesterases (e.g. acid phosphatases and phytases) and a shift in functional ability in these soils towards nutrient acquisition from the organic P pool (Figures 5 and 6).

At Haast, both surveyed phosphatase genes, (class A) and (*phoD*), appeared to respond similarly to the prevailing conditions of the dunes along with significant taxonomic shifts (Figure 7; Supplemental Figure S6). Soil nutrient levels (e.g. total P) and stoichiometry (e.g. C:N, C:P, N:P) were found to be predictors of community diversity and composition or abundance of class A acid phosphatase communities (Luo *et al.* 2019; Zheng *et al.* 2019; Gaiero *et al.* 2020). *phoD*-harbouring communities may also respond uniquely to resource limitations and available P resources (Jorquera *et al.* 2014; Margalef *et al.* 2017; Gaiero *et al.* 2020). Admittedly, the abundance and diversity of bacterial phosphatases and their drivers is complex (Spohn and Kuzyakov 2013; Margalef *et al.* 2017; Neal *et al.* 2018; Gaiero *et al.* 2020), and there is still a need to better understand their biogeography and ecological roles (George *et al.* 2018). Of the

known contributors to these genes, Proteobacteria are often overrepresented in genome databases (Ragot, Kertesz and Bünemann 2015; Gaiero *et al.* 2018). The results at Haast also indicated a narrow phylogenetic range, but with caveats in ascribing taxonomic identities to these functional genes. However, many of the changes observed at the order level for both class A and *phoD* were also coincident with changes in the wider bacterial community composition (Supplemental Figures S3, S6), and in the taxonomic contribution to predicted P-cycling enzymes (Supplemental Figure S5). This can be seen in many of the dominant phosphatase-gene harbouring taxa, Rhizobiales, Pseudomonadales, Xanthomonadales, Sphingomonadales, Syntrophobacterales (Supplemental Figures S3, S6).

Positive trends found between quantified acid and alkaline phosphatases across the chronosequence were of interest and not entirely expected in these low pH soils (Lidbury *et al.* 2017). *phoD* genes showed the same trends in PICRUSt2 (K01113) as qPCR, but the changes were not significant as compared to quantitation (Figures 5 and 6). Further, the unchanging alkaline phosphatase EC abundance (EC 3.1.3.1) can be explained by alkaline phosphatase genes *phoA* and *phoB* (Figure 6). Predominantly found in Enterobacteria (Lin *et al.* 2015), they appear to be contributing meaningfully to the EC function and tended to be higher in the young dunes (Figure 6).

We have also explored phytase genes (e.g. BPP) using the primers available and with predictive metagenomics (PICRUSt2). These genes should continue to be targeted in the future and when more molecular tools are available. Metagenomic analyses have shown that phytases, particularly 3-phytase, respond strongly to phosphorus depletion in long-term tropical forest fertilization experiments (Yao *et al.* 2018). Further, the genetic potential of phosphonate metabolism remained relatively consistent throughout the chronosequence, with the exception of

an increased abundance of *phnP* (EC 3.1.4.55), part of the greater CP-lyase *phn* operon (Figures 5 and 6). Previous work revealed that phosphonates remained low in abundance (1-3% of total organic P) and declined to undetectable levels after 3384 years (dune 11) (Turner, Wells and Condron 2014). Molecular analyses suggest these phosphonate compounds may be more abundant but are turning over rapidly throughout the chronosequence. Finally, we would be remiss not to mention the contributions to P-cycling by fungi as well. Fungi, such as *Aspergillus*, *Penicillium* and *Trichoderma*, are well known to produce phytases, phosphomonoesterases, and contribute to P solubilizing activities (Kapri and Tewari 2010; Richardson and Simpson 2011). As we have seen at Haast, fungi are increasingly important as the soils age, in agreement with our initial hypothesis. Without predictive tools for fungi, such as PICRUSt2 for bacteria, we might glean functional insights at a lower taxonomic level. For fungi, we were unable to find significant taxonomic differences below the order level, as we found for Hypocreales in the older dunes (Ascomycota). However, an elevated level of *Trichoderma* spp. (of the order Hypocreales) in the older dunes does suggest a great potential for P cycling (Kapri and Tewari 2010).

Predictive algorithms such as PICRUSt2 have opened new lines of research to explore but these methods have certain limitations, as do other metagenomic techniques (Douglas *et al.* 2020). Since gene predictions are based on existing reference genomes, there are limitations for most sample types, and more so with samples containing highly complex communities such as soils (Douglas *et al.* 2020). As such, this approach is less capable in terms of rare or strain-specific functions (Douglas *et al.* 2020). Nonetheless, we found empirical support for PICRUSt2 predictions, namely in non-specific acid phosphatase class A gene quantitation, trends in acid phosphatase activity (Table 1), and historical phosphomonoesterase and phosphodiesterase activities (Figure 1D). The PICRUSt2 predictions will enable subsequent efforts, in targeting

specific genes by PCR approaches, and by metagenomic analyses which would be novel to this chronosequence.

5. Conclusions

This chronosequence study allowed us to characterize the dynamics of soil microbial communities in the upper mineral layer as they relate to and participate in ecosystem development. Our findings revealed that bacterial diversity increased with ecosystem development due to declining abundance of predominant Firmicutes and Proteobacteria in the older dunes. Highest fungal diversity occurred after 392 years, coincident with maximum tree basal area and organic P accumulation. The Bacteria:Fungi ratio decreased as soils developed, primarily resulting from a decline in bacterial abundance. However, network analysis revealed a continued dominance of bacterial taxa, amid increasingly competitive and interconnected soil microbial communities. There was evidence of a shift in microbial functional ability over time, with increasing potential for organic P mineralization as P availability declined during pedogenesis and ecosystem succession. Functional inferences from metagenome predictions supported the empirical observations that acid phosphatases and phytases were particularly important in the older soils of the chronosequence, and revealed gene targets to explore in future work.

6. Acknowledgements

This research was supported by funding provided by Lincoln University (L.M. Condron), Canada Research Chairs Program, and an NSERC Discovery Grant (K.E. Dunfield). Special thanks to Andrew Wells for soil sampling. We greatly appreciate discussions and advice provided by colleagues Dasiel Obregón Alvarez and Eduardo Kovalski Mitter.

7. References

Afgan E, Baker D, Batut B *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 2018;**46**:W537–44.

Allison VJ, Condron LM, Peltzer DA *et al.* Changes in enzyme activities and soil microbial community composition along carbon and nutrient gradients at the Franz Josef chronosequence, New Zealand. *Soil Biol Biochem* 2007;**39**:1770–81.

Altschul S. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;**25**:3389–402.

Augusto L, Achat DL, Jonard M *et al.* Soil parent material—A major driver of plant nutrient limitations in terrestrial ecosystems. *Glob Chang Biol* 2017;**23**:3808–24.

Bahram M, Anslan S, Hildebrand F *et al.* Newly designed 16S rRNA metabarcoding primers amplify diverse and novel archaeal taxa from the environment. *Environ Microbiol Rep* 2019;**11**:487–94.

Banerjee S, Kirkby CA, Schmutter D *et al.* Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. *Soil Biol Biochem* 2016;**97**:188–98.

Barbera P, Kozlov AM, Czech L *et al.* EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. Posada D (ed.). *Syst Biol* 2019;**68**:365–9.

Bastian M, Heymann S, Jacomy M. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Third Int AAAI Conf Weblogs Soc Media* 2009:361–2.

Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B* 1995;**57**:289–300.

Bergkemper F, Kublik S, Lang F *et al.* Novel oligonucleotide primers reveal a high diversity of

microbes which drive phosphorous turnover in soil. *J Microbiol Methods* 2016a;**125**:91–7.

Bergkemper F, Schöler A, Engel M *et al.* Phosphorus depletion in forest soils shapes bacterial communities towards phosphorus recycling systems. *Environ Microbiol* 2016b;**18**:1988–2000.

Bergmann GT, Bates ST, Eilers KG *et al.* The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biol Biochem* 2011;**43**:1450–5.

Bokulich NA, Kaehler BD, Rideout JR *et al.* Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 2018;**6**:90.

Bolyen E, Rideout JR, Dillon MR *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;**37**:852–7.

Callahan BJ, McMurdie PJ, Rosen MJ *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;**13**:581–3.

Caspi R, Billington R, Fulcher CA *et al.* The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res* 2018;**46**:D633–9.

Crowther TW, Boddy L, Hefin Jones T. Functional and ecological consequences of saprotrophic fungus–grazer interactions. *ISME J* 2012;**6**:1992–2001.

Czech L, Barbera P, Stamatakis A. Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics* 2020;**36**:3263–5.

Dereeper a, Guignon V, Blanc G *et al.* Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 2008;**36**:W465-9.

Douglas GM, Maffei VJ, Zaneveld JR *et al.* PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* 2020;**38**:685–8.

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7.

Eger A, Almond PC, Condron LM. Pedogenesis, soil mass balance, phosphorus dynamics and vegetation communities across a Holocene soil chronosequence in a super-humid climate, South Westland, New Zealand. *Geoderma* 2011;**163**:185–96.

Feng K, Zhang Y, He Z *et al.* Interdomain ecological networks between plants and microbes. *Mol Ecol Resour* 2019;**19**:1565–77.

Fernandes AD, Macklaim JM, Linn TG *et al.* ANOVA-Like Differential Expression (ALDEx) Analysis for Mixed Population RNA-Seq. Parkinson J (ed.). *PLoS One* 2013;**8**:e67019.

Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 2017;**15**:579–90.

Fraser TD, Lynch DH, Gaiero J *et al.* Quantification of bacterial non-specific acid (*phoC*) and alkaline (*phoD*) phosphatase genes in bulk and rhizosphere soil from organically managed soybean fields. *Appl Soil Ecol* 2017;**111**:48–56.

Friedman J, Alm EJ. Inferring Correlation Networks from Genomic Survey Data. *PLoS Comput Biol* 2012;**8**:1–11.

Fuerst JA, Sagulenko E. Beyond the bacterium: planctomycetes challenge our concepts of microbial structure and function. *Nat Rev Microbiol* 2011;**9**:403–13.

Gaiero JR, Bent E, Boitt G *et al.* Effect of long-term plant biomass management on phosphatase-producing bacterial populations in soils under temperate grassland. *Appl Soil Ecol* 2020;**151**:103583.

Gaiero JR, Bent E, Fraser TD *et al.* Validating novel oligonucleotide primers targeting three classes of bacterial non-specific acid phosphatase genes in grassland soils. *Plant Soil*

2018;**427**:39–51.

George TS, Giles CD, Menezes-Blackburn D *et al.* Organic phosphorus in the terrestrial environment: a perspective on the state of the art and future priorities. *Plant Soil* 2018;**427**:191–208.

Gilbert JA, Thomas S, Cooley NA *et al.* Potential for phosphonoacetate utilization by marine bacteria in temperate coastal waters. *Environ Microbiol* 2009;**11**:111–25.

Hammer Ø, Harper DAT a. T, Ryan PD. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontol Electron* 2001;**4**:1–9.

Hothorn T, Bretz F, Westfall P. Simultaneous Inference in General Parametric Models. *Biometrical J* 2008;**50**:346–63.

Huang H, Shi P, Wang Y *et al.* Diversity of beta-propeller phytase genes in the intestinal contents of grass carp provides insight into the release of major phosphorus from phytate in nature. *Appl Environ Microbiol* 2009;**75**:1508–16.

Jangid K, Whitman WB, Condron LM *et al.* Progressive and retrogressive ecosystem development coincide with soil bacterial community change in a dune system under lowland temperate rainforest in New Zealand. *Plant Soil* 2013a;**367**:235–47.

Jangid K, Whitman WB, Condron LM *et al.* Soil bacterial community succession during long-term ecosystem development. *Mol Ecol* 2013b;**22**:3415–24.

Jorquera MA, Martínez OA, Marileo LG *et al.* Effect of nitrogen and phosphorus fertilization on the composition of rhizobacterial communities of two Chilean Andisol pastures. *World J Microbiol Biotechnol* 2014;**30**:99–107.

Kapri A, Tewari L. Phosphate solubilization potential and phosphatase activity of rhizospheric Trichoderma spp. *Brazilian J Microbiol* 2010;**41**:787–95.

Katoh K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002;**30**:3059–66.

Kielak AM, Barreto CC, Kowalchuk GA *et al.* The Ecology of Acidobacteria: Moving beyond Genes and Genomes. *Front Microbiol* 2016;**7**, DOI: 10.3389/fmicb.2016.00744.

Kielak AM, Castellane TCL, Campanharo JC *et al.* Characterization of novel Acidobacteria exopolysaccharides with potential industrial and ecological applications. *Sci Rep* 2017;**7**:41193.

Kramer S, Dibbern D, Moll J *et al.* Resource Partitioning between Bacteria, Fungi, and Protists in the Detritusphere of an Agricultural Soil. *Front Microbiol* 2016;**7**:1524.

Laliberté E, Grace JB, Huston MA *et al.* How does pedogenesis drive plant diversity? *Trends Ecol Evol* 2013;**28**:331–40.

Letunic I, Bork P, D.H. H *et al.* Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 2016;**44**:W242–5.

Lidbury IDEA, Fraser T, Murphy ARJ *et al.* The 'known' genetic potential for microbial communities to degrade organic phosphorus is reduced in low-pH soils. *Microbiologyopen* 2017;**6**:e00474.

Lin X, Wang L, Shi X *et al.* Rapidly diverging evolution of an atypical alkaline phosphatase (PhoAaty) in marine phytoplankton: Insights from dinoflagellate alkaline phosphatases. *Front Microbiol* 2015;**6**:1–12.

Louca S, Doebeli M. Efficient comparative phylogenetics on large trees. *Bioinformatics* 2018;**34**:1053–5.

Luo G, Sun B, Li L *et al.* Understanding how long-term organic amendments increase soil phosphatase activities: Insight into phoD- and phoC-harboring functional microbial

populations. *Soil Biol Biochem* 2019;**139**:107632.

Margalef O, Sardans J, Fernández-Martínez M *et al.* Global patterns of phosphatase activity in natural soils. *Sci Rep* 2017;**7**:1337.

Neal AL, Blackwell M, Akkari E *et al.* Phylogenetic distribution, biogeography and the effects of land management upon bacterial non-specific Acid phosphatase Gene diversity and abundance. *Plant Soil* 2018;**427**:175–89.

Nguyen NH, Song Z, Bates ST *et al.* FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol* 2016;**20**:241–8.

Nilsson RH, Larsson K-H, Taylor AFS *et al.* The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res* 2019;**47**:D259–64.

Olsen S. *Estimation of Available Phosphorus in Soils by Extraction With Sodium Bicarbonate*. Washington, D.C.: United States Department of Agriculture, 1954.

Peltzer D a, Wardle DA, Allison VJ *et al.* Understanding ecosystem retrogression. *Ecol Monogr* 2010;**80**:509–29.

Price MN, Dehal PS, Arkin AP. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. Poon AFY (ed.). *PLoS One* 2010;**5**:e9490.

Quast C, Pruesse E, Yilmaz P *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**:590–6.

R Core Team. R: A language and environment for statistical computing. 2020.

Ragot SA, Kertesz MA, Bünemann EK. phoD Alkaline Phosphatase Gene Diversity in Soil. *Appl Environ Microbiol* 2015;**81**:7281–9.

Reji L, Tolar BB, Smith JM *et al.* Differential co-occurrence relationships shaping ecotype

diversification within Thaumarchaeota populations in the coastal ocean water column. *ISME J* 2019;**13**:1144–58.

Richardson AE, Simpson RJ. Soil Microorganisms Mediating Phosphorus Availability Update on Microbial Phosphorus. *Plant Physiol* 2011;**156**:989–96.

Roberts K, Defforey D, Turner BL *et al.* Oxygen isotopes of phosphate and soil phosphorus cycling across a 6500year chronosequence under lowland temperate rainforest. *Geoderma* 2015;**257**–**258**:14–21.

Smilauer P, Lepš J. *Multivariate Analysis of Ecological Data Using CANOCO 5*. Cambridge: Cambridge University Press, 2014.

Spohn M, Kuzyakov Y. Phosphorus mineralization can be driven by microbial need for carbon. *Soil Biol Biochem* 2013;**61**:69–75.

Tabatabai M a., Bremner JM. Use of p-nitrophenyl phosphate for assay of soil phosphatase activity. *Soil Biol Biochem* 1969;**1**:301–7.

Taiyun Wei M, Taiyun Wei cre A, Simko aut V *et al. Package "corrplot" Title Visualization of a Correlation Matrix*., 2017.

Tamburini F, Pfahler V, Bünemann EK *et al.* Oxygen Isotopes Unravel the Role of Microorganisms in Phosphate Cycling in Soils. *Environ Sci Technol* 2012;**46**:5956–62.

Tan H, Barret M, Mooij MJ *et al.* Long-term phosphorus fertilisation increased the diversity of the total bacterial community and the phoD phosphorus mineraliser group in pasture soils. *Biol Fertil Soils* 2013;**49**:661–72.

Turner BL, Condron LM, Wells A *et al.* Soil nutrient dynamics during podzol development under lowland temperate rain forest in New Zealand. *Catena* 2012a;**97**:50–62.

Turner BL, Lambers H, Condron LM *et al.* Soil microbial biomass and the fate of phosphorus

during long-term ecosystem development. *Plant Soil* 2013;**367**:225–34.

Turner BL, Wells A, Andersen KM *et al.* Patterns of tree community composition along a coastal dune chronosequence in lowland temperate rain forest in New Zealand. *Plant Ecol* 2012b;**213**:1525–41.

Turner BL, Wells A, Condron LM. Soil organic phosphorus transformations along a coastal dune chronosequence under New Zealand temperate rain forest. *Biogeochemistry* 2014;**121**:595–611.

Uksa M, Schloter M, Endesfelder D *et al.* Prokaryotes in Subsoil—Evidence for a Strong Spatial Separation of Different Phyla by Analysing Co-occurrence Networks. *Front Microbiol* 2015;**6**:1269.

Vick-Majors TJ, Priscu JC, A Amaral-Zettler L. Modular community structure suggests metabolic plasticity during the transition to polar night in ice-covered Antarctic lakes. *ISME J* 2014;**8**:778–89.

Walker TW, Syers JK. The fate of phosphorus during pedogenesis. *Geoderma* 1976;**15**:1–19.

Wardle DA, Jonsson M, Bansal S *et al.* Linking vegetation change, carbon sequestration and biodiversity: Insights from island ecosystems in a long-term natural experiment. *J Ecol* 2012;**100**:16–30.

Wells A, Goff J. Coastal dunes in Westland, New Zealand, provide a record of paleoseismic activity on the Alpine fault. *Geology* 2007;**35**:731–4.

Williams MA, Jangid K, Shanmugam SG *et al.* Bacterial communities in soil mimic patterns of vegetative succession and ecosystem climax but are resilient to change between seasons. *Soil Biol Biochem* 2013;**57**:749–57.

Williamson WM, Wardle DA, Yeates GW. Changes in soil microbial and nematode

communities during ecosystem decline across a long-term chronosequence. *Soil Biol Biochem* 2005;**37**:1289–301.

Yao Q, Li Z, Song Y *et al.* Community proteogenomics reveals the systemic impact of phosphorus availability on microbial functions in tropical soil. *Nat Ecol Evol* 2018;**2**:499–509.

Ye Y, Doak TG. A parsimony approach to biological pathway reconstruction/inference for genomes and metagenomes. Ouzounis CA (ed.). *PLoS Comput Biol* 2009;**5**:e1000465.

Zheng MM, Wang C, Li WX *et al.* Soil Nutrients Drive Function and Composition of phoC-Harboring Bacterial Community in Acidic Soils of Southern China. *Front Microbiol* 2019;**10**:2654.

Zhou J, Xia B, Treves DS *et al.* Spatial and Resource Factors Influencing High Microbial Diversity in Soil. *Appl Environ Microbiol* 2002;**68**:326–34.

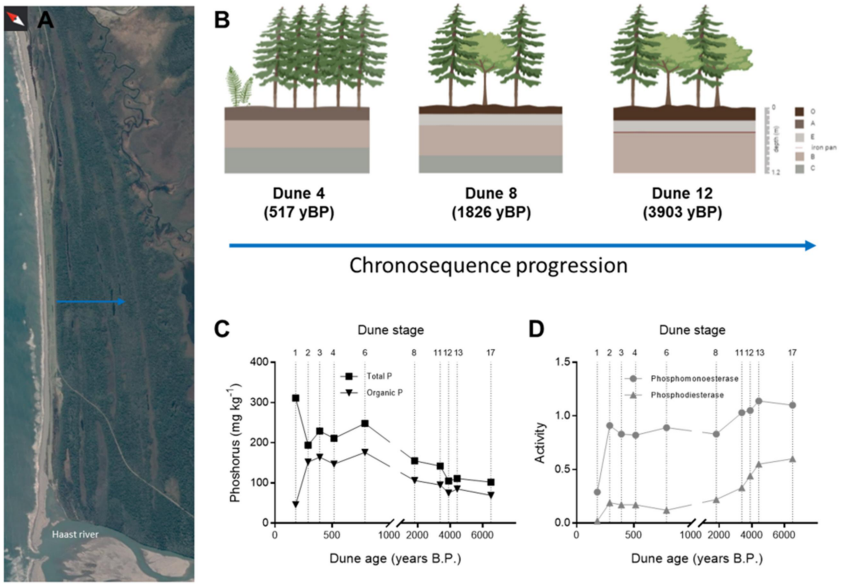


Figure 1. Overview of the Haast dune chronosequence. A) Aerial view of the chronosequence and sampling transect (blue); B) forest community and soil profile to 1.2 m depth at three stages of succession (forest vegetation included ferns, conifers such as *Dacrydium cupressinum* and/or woody angiosperms like *Weinmannia racemosa*) (Turner *et al.* 2012b); C) historical P data (Turner *et al.* 2012a; Turner, Wells and Condron 2014); D) historical phosphatase activity measured from mineral soils (0-20 cm) using the 4-methylumbelliferone-linked fluorogenic substrates (unpublished).

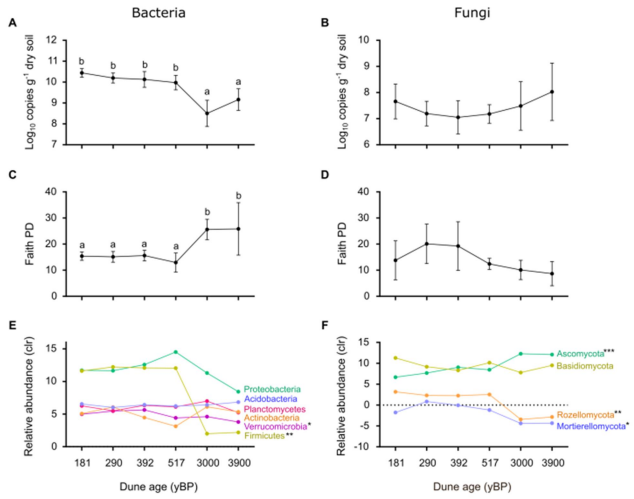


Figure 2. Abundance and diversity of soil bacterial (A, C, E) and fungal communities (B, D, F). A, B) Total bacterial (16S) and fungal (18S) rRNA gene copies; C, D) Faith's phylogenetic diversity; E, F) Relative abundance (mean centered log-ratio, clr) of dominant phyla. Letter-groups denote significant pairwise differences in gene copies (A, B) or Faith's phylogenetic diversity (C, D) between dunes (Tukey's HSD test). Relative abundance of phyla (E, F) was calculated as clr and compared between young (181-290 yBP) and old dunes (3000-3900 yBP) by Wilcoxon test with Benjamini-Hochberg FDR correction ('ALDEx2' in R). Sensitive phyla are denoted by asterisks (· $P < 0.1$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$).

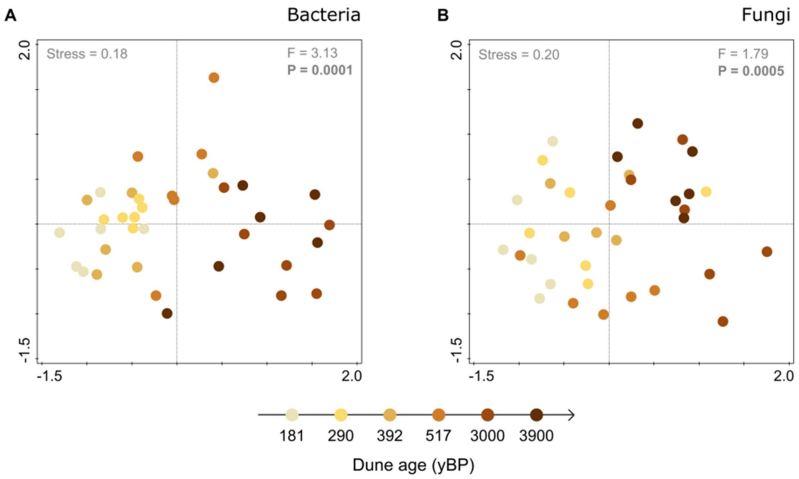
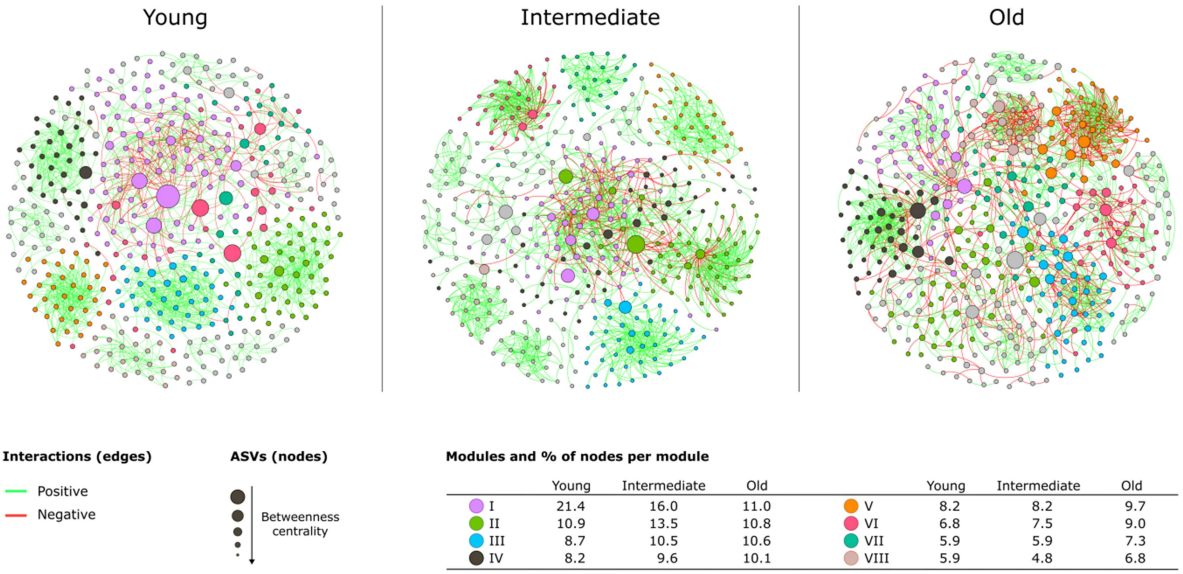


Figure 3. Non-metric multidimensional scaling (NMDS) ordination plots of Bray-Curtis dissimilarities (k = 2, stress values shown) for soil A) bacterial and B) fungal communities. Changes in community composition between dunes were tested using PERMANOVA (F and P values shown). The transition of dune colours from light to dark represent the progression of soil age (years before present, yBP) from 181 to 3900 yBP.



	Young	Intermediate	Old		Young	Intermediate	Old
I	21.4	16.0	11.0	V	8.2	8.2	9.7
II	10.9	13.5	10.8	VI	6.8	7.5	9.0
III	8.7	10.5	10.6	VII	5.9	5.9	7.3
IV	8.2	9.6	10.1	VIII	5.9	4.8	6.8

Figure 4. Co-occurrence networks of soil bacterial and fungal communities from dunes at three designated periods of the Haast chronosequence: 'young' (181-290 yBP), 'intermediate' (392-517 yBP), and 'old' dunes (3000-3900 yBP). Negative (red) and positive (green) interactions (edges) were calculated using SparCC ($P < 0.01$, $R > 0.6$ or < -0.6) of the 500 most abundant ASVs (nodes). The size of the nodes represents betweenness centrality, and the top 8 modules, ranked by percentage of nodes per module, are shown by coloured nodes.

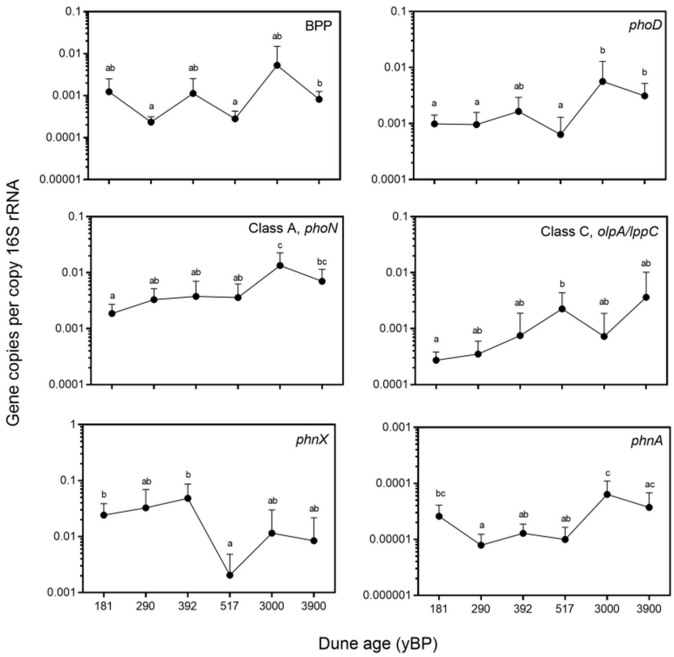


Figure 5. Relative abundance of bacterial genes related to organic P mineralization (copies per unit 16S rRNA) found in mineral soils along the Haast chronosequence. Pairwise significance between dunes is shown by letter-group (Tukey's HSD test), with most genes having $P < 0.01$, except for BPP ($P < 0.05$) and class C ($P = 0.1$). Absolute gene and transcript abundance (copies g^{-1} dry soil) can be found in Figure S4.

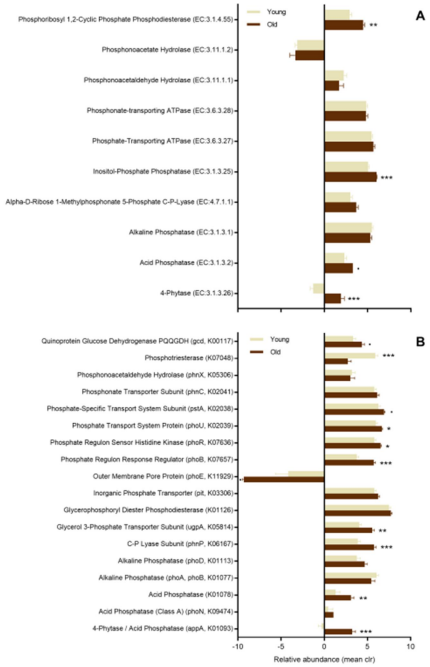


Figure 6. Predicted abundance of P-cycling enzymes (A) (EC: enzyme commission) and genes (KOs: Kegg orthologs) obtained using PICRUSt2. Relative abundances (centered log-ratio, clr) were compared between young (181-290 yBP; dunes 1 and 2) and old dunes (3000-3900 yBP; dunes 10 and 12) to track the major changes taking place through the Haast chronosequence. Statistical comparisons were made using Wilcoxon Rank Sum test with Benjamini-Hochberg FDR correction ('ALDEx2' in R). Significant differences between young and old dunes are denoted by asterisks ($P < 0.1$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$).

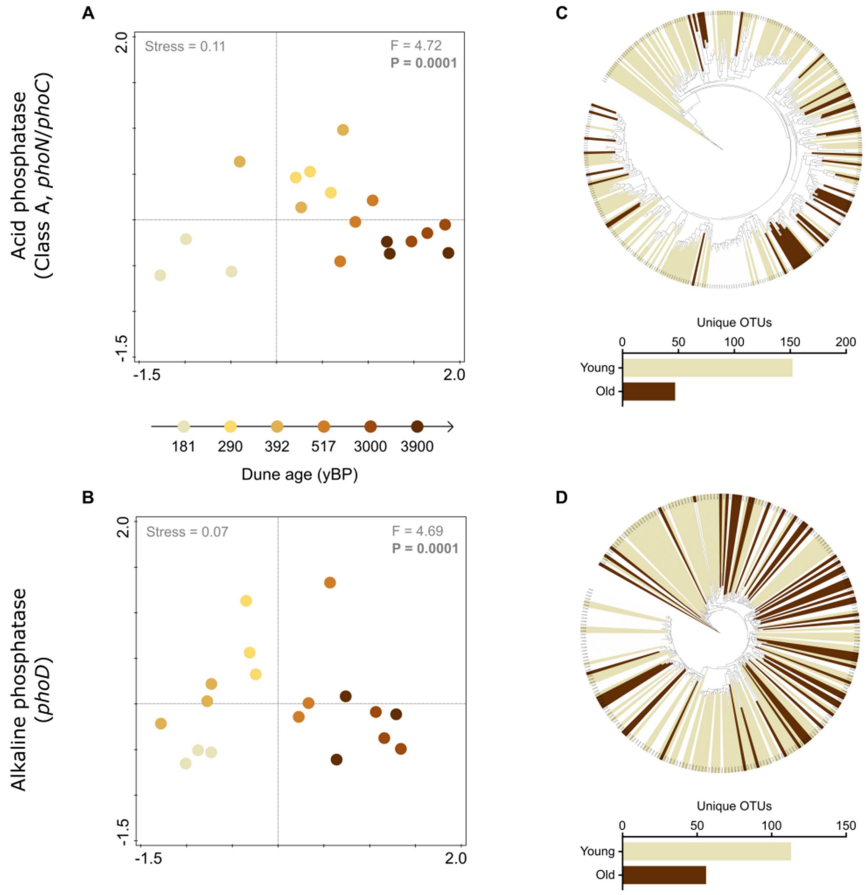


Figure 7. Compositional changes in soil bacterial non-specific acid phosphatase (class A, *phoN*/*phoC*) (A, C) and alkaline phosphatase (*phoD*) (B, D) genes through the Haast chronosequence to 3900 yBP. A, B) Non-metric MDS ordination plots of Bray-Curtis dissimilarities (k = 2, stress values shown); C, D) phylogenetic trees and barplots of class A and *phoD* OTUs with unique OTUs (coloured) found in young (181-290 yBP) and old (3000-3900 yBP) dunes. Changes in community composition between young and old dunes were tested using PERMANOVA (F and P values shown in A and B).

Table 1. Chemical and biological properties of mineral soils found across the Haast chronosequence. Mean values are presented per dune, with standard deviation in parentheses, and results of significance testing[a].

Dune system	1	2	3	4	10	12
Dune age (yBP)	181	290	392	517	3000	3900
Organic horizon	<1	5	6	3	-	22
depth (cm)[b]	-	(0.4)	(0.7)	(0.5)	-	(1)
pH	4.65	4.52	4.62	4.08	4.00	4.08
	(0.44)	(0.13)	(0.45)	(0.17)	(0.13)	(0.08)
	a	a	ab	bc	c	c
Total C (%)	3.31	3.05	4.07	5.09	8.15	11.41
	(1.08)	(0.54)	(0.88)	(1.77)	(2.69)	(8.12)
	ab	a	abc	bd	d	cd
Total N (%)	0.25	0.19	0.23	0.28	0.36	0.54
	(0.06)	(0.03)	(0.04)	(0.08)	(0.14)	(0.39)
	ab	a	ab	ab	b	ab
C:N	13.18	15.78	17.45	18.05	23.75	21.47
	(1.20)	(0.80)	(1.60)	(1.67)	(4.44)	(2.79)
	a	b	bc	cd	e	de
Olsen P (mg P kg^{-1})	5.33	5.83	5.33	5.17	4.67	6.83
	(0.82)	(0.75)	(1.51)	(0.75)	(1.75)	(4.22)
	ns	ns	ns	ns	ns	ns
Acid phosphatase activity	5.71	3.49	5.28	6.51	7.07	10.72
(μmol ρ-NP g^{-1} h^{-1})	(1.86)	(0.85)	(1.31)	(4.57)	(4.62)	(8.52)
	ns	ns	ns	ns	ns	ns
Bacteria:Fungi ratio[c]	1528	1709	2758	1118	17	164
	(2113)	(1749)	(2564)	(1238)	(22)	(272)
	bc	c	c	c	a	ab

[a] Letters indicate differences between dunes ($P < 0.05$, Tukey's HSD)

[b] Above the sampled horizon. From (Turner *et al.* 2012a)

[c] based on qPCR results for 16S and 18S rRNA gene copies

Table 2. Topological properties of the microbial co-occurrence networks[a]. Young dunes (181-290 yBP), intermediate dunes (392-517 yBP), old dunes (3000-3900 yBP).

Network metrics	Young	Intermediate	Old
Nodes	439	438	455
Edges	1381	1591	1582
Modularity (no. of modules)	0.841 (24)	0.833 (19)	0.782 (17)
Density	0.014	0.017	0.015
Average degree	6.292	7.265	6.954
Clustering coefficient	0.467	0.5	0.374
Average path length	5.643	5.203	4.896
Edges (positive and negative)			
B-B	388 (28.1%)	444 (27.9%)	853 (53.9%)
B-F	209 (15.1%)	299 (18.8%)	485 (30.7%)
F-F	784 (56.8%)	848 (53.3%)	244 (15.4%)
Edges (negative)			
Total	201 (14.6%)	245 (15.4%)	582 (36.8%)
B-B	167 (83.1%)	159 (64.9%)	326 (56.0%)
B-F	34 (16.9%)	85 (34.7%)	256 (44.0%)
F-F	0 (0.0%)	1 (0.4%)	0 (0.0%)

[a] Co-occurrence networks were generated using SparCC correlation results of the 500 most abundant fungal and bacterial ASVs. Networks included positive and negative correlations, using a cutoff of $-0.6 > R > 0.6$ ($P < 0.01$). A total of 12 samples were included in each network.