Soil Biology and Biochemistry 135 (2019) 163–172

Contents lists available at ScienceDirect

Soil Biology and Biochemistry

journal homepage: www.elsevier.com/locate/soilbio

Changes in microbial functional genes within the soil metagenome during forest ecosystem restoration

Shan Sun*, Brian D. Badgley

School of Plant and Environmental Sciences, Virginia Tech, Blacksburg, VA, 24060, USA

ARTICLE INFO

Keywords:
Microbial community
Metagenome
Microbial functional gene
Forest ecosystem restoration

ABSTRACT

As important mediators of numerous ecological processes, soil microorganisms play essential roles in the recovery of ecosystems after disturbance. Using next-generation sequencing techniques, microbial taxonomic changes during ecosystem restoration have been widely studied, but data describing microbial community structure alone can be difficult to link to ecosystem processes mediated by microorganisms. Shotgun metagenome sequencing provides a chance to examine changes among thousands of functional genes during the recovery of microbial communities and ecological function. We analyzed 15 soil metagenomes from a chronosequence of mine soils spanning 6–31 years since reforestation along with unmined reference soils. Taxonomic and functional changes indicate a shift from copiotrophic to oligotrophic groups, increasing metabolism of recalcitrant carbon sources and the influence of vegetation. Increases in genes involved in transposable elements, virulence, defense, and stress response suggest more cooperative and competitive interactions among microorganisms with chronosequence age. Within N cycling groups, ammonia and nitrite oxidizing bacteria increased significantly during restoration, but few significant changes were observed in key N-cycle functional genes. The low relative abundances of methanotrophs and methane monooxygenase genes in all reforested soils explains previous observations that methane consumption has not recovered at these sites 31 years after reforestation. This work helps identify possible mechanisms linking the soil microbiome to ecosystem recovery, with a specific focus on N cycling and greenhouse gas emission, to better understand the roles soil microorganisms play in the restoration of ecosystem functions.

1. Introduction

Forest ecosystems provide important ecological services, playing essential roles in global biogeochemical cycling and regulation of atmospheric greenhouse gases (Bonan, 2008). Within forests, soil supports most ecosystem processes (Baldrian et al., 2012) and many are mediated by soil microbial communities (Baldrian et al., 2012; Song et al., 2016). Soil microorganisms play important roles in C and N cycling, which influence ecosystem services such as C sequestration and N retention (Six et al., 2006; De Vries et al., 2011). For example, the largest biological sinks for atmospheric methane are the methanotrophs living in forest soils (Kolb, 2009). Therefore, knowledge about soil microbial communities and their interactions with the environment following disturbance is important for understanding the recovery patterns of ecosystem function.

Microbial communities are responsive to varying environmental factors and are likely under the regulation of available resources, vegetation inputs, environmental stress, and biological interactions (Paver et al., 2013; Knelman et al., 2014; Zhou et al., 2014). The influence of plant litter deposition and root exudates on microbial communities suggests that vegetation could be an important driver of microbial succession after disturbances (Zumsteg et al., 2012; Salles et al., 2017), and changes in resource availability can also influence microbial succession (Knelman et al., 2014). Previous studies have observed increases in microbial biomass, respiration, N mineralization and nitrification rates during succession (Zak et al., 1990; Yan et al., 2009; Cline and Zak, 2015). Studies of microbial taxonomic changes during ecosystem succession suggest that microbial communities follow gradual changes with increasing age, which likely reflect corresponding changes in environmental conditions but also probably impact the recovery of ecosystem functions (Banning et al., 2011; Dini-Andreote et al., 2014; Sun et al., 2017).

Describing only taxonomic changes in soil microbiota can provide limited information regarding function, which then has to be inferred based upon the assumption of ecological coherence within taxonomic groups (Philippot et al., 2010; Fierer et al., 2012b). In contrast, changes

* Corresponding author. School of Plant and Environmental Sciences Virginia Tech, 1880 Pratt Drive, Room 1004 Blacksburg, Virginia, 24060, USA.
E-mail address: shansun@vt.edu (S. Sun).

https://doi.org/10.1016/j.soilbio.2019.05.004
Received 30 November 2018; Received in revised form 3 May 2019; Accepted 6 May 2019
Available online 06 May 2019
0038-0717/

in the abundances of specific functional genes within the soil microbial community during recovery can provide more direct insight into how environmental changes impact microbial processes that relate to ecosystem functions. For example, the nitrogen cycle is of particular importance in forest ecosystems, where N is usually a major limiting nutrient for plants and microorganisms (Akselsson and Westling, 2005). Forest clearing and disturbance can disrupt N transformations (Bukata and Kyser, 2005), so the return of normal nitrogen cycling during reforestation is essential for the recovery of forest ecosystems. Previous studies focused on nitrogen related functional genes relied on quantitative polymerase chain reaction (qPCR) to quantify individual genes of interest from mixed microbial communities in soil samples. Abundances of genes involved in N fixation (*nifH*) and denitrification (*nirS*, *nirK*) in a salt marsh peaked at intermediate succession (35 yr) (Salles et al., 2017), while studies in deglaciated forelands have shown that denitrification (*nirS*) and ammonia oxidation (ammonia oxidizing bacteria-*amoA*) can either decrease (Brankatschk et al., 2011) or increase (Zeng et al., 2016) with succession age. Beyond the N cycling, forests can also serve as a significant sink for greenhouse gases such as CO_2, CH_4 (Böttcher et al., 2012), and N_2O under certain conditions (Chapuis-Lardy et al., 2007). However, land use change and deforestation can contribute significantly to global greenhouse gases emissions (Fearnside, 2000), and the recovery of greenhouse gas uptake may take > 100 years to recover (Smith et al., 2000).

Advances in shotgun DNA sequencing methods in recent years have greatly increased the capability to characterize environmental metagenomes, which provide an extensive view of the relative abundances of thousands of functional genes present in an environmental sample. This is in contrast to targeting single genes individually with qPCR, which can be further complicated by functional redundancy across diverse phylogenetic groups (Wei et al., 2015; Fullerton et al., 2017). However, metagenomics has only been used in a few studies of ecosystem restoration and succession. These include changes in functional genes between pre-agricultural tallgrass prairie and modern agricultural soils (Fierer et al., 2013), the impact of engineered soil formulations on microbial functions in restored mine sites (Kumaresan et al., 2017), and how plant-driven changes shape microbial communities during succession post agricultural abandonment (Cline and Zak, 2015).

In this work, our goal was to expand on that understanding by using metagenomics to track changes in microbial function during multiple stages of reforestation and ecosystem restoration. We revisited a chronosequence of reforested reclaimed mine sites spanning 6–31 years plus unmined reference sites where we have previously described taxonomic changes in the soil microbiota (Sun et al., 2017). We used full shotgun metagenomics to directly measure changes in functional gene abundance among the soil microbiomes during ecosystem recovery. We also investigated whether changes in known N and CH_4 cycling taxa or genes were related to changes in relevant ecosystem processes rates measured previously at these same sites by Avera et al. (2015). Finally, we identified other changes in the soil microbiome related to biotic interactions, which are hypothesized to increase over time along this chronosequence based on previous network analysis (Sun et al., 2017).

2. Materials and methods

2.1. Sampling and DNA extraction

The soil samples for this study were collected in September 2014 from the same chronosequence of reforested mined sites as described in previous work (Avera et al., 2015; Sun et al., 2017). Briefly, the chronosequence was located in Wise County, VA, USA, and included sites that had been recovering for 6, 12, 22, and 31 yr since reforestation at the time of sampling, along with a nearby unmined reference forest site as a control. Triplicate circular plots (10 m diameter) were established at each site, controlling for landscape and vegetation differences; details of plot design were described by Avera et al. (2015). For this study, three 5 cm * 5 cm soil samples from each plot were excavated by trowel from 0 to 10 cm depth. The three soil samples were then pooled and homogenized to better capture within-plot heterogeneity. Composited samples of three plots from five different chronosequence sites (ages 6, 12, 22, 31 and unmined reference) resulted in 15 total metagenomes. The samples were transported to the lab on ice and stored at −80 °C until DNA was extracted from approximately 0.25 g soil using the PowerSoil Soil DNA isolation kit (MoBio Laboratories, Carlsbad, CA, USA). DNA concentrations were measured with Qubit 2.0 fluorometer (Invitrogen, USA).

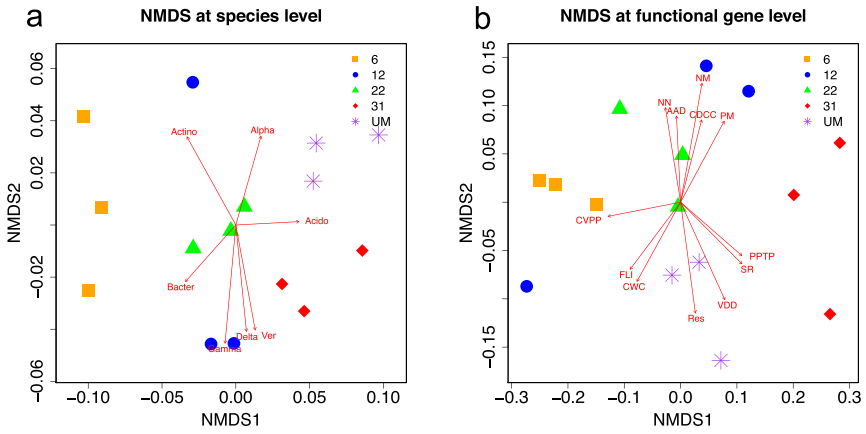
2.2. Shotgun metagenome sequencing and sequence processing

For shotgun metagenomics, DNA concentrations from all samples were equalized with PCR water (MoBio Laboratories, Carlsbad, CA, USA) and sheared to 300 bp using an M220 Focused-ultrasonicator (Covaris, USA) in 50 μL tubes for 90 s. All 15 DNA libraries were then multiplexed and sequenced on the Illumina Hiseq platform in a 101 bp paired-end sequencing run at the Virginia Tech Biocomplexity Institute. The raw sequence reads were uploaded to MG-RAST (Meyer et al., 2008) for quality control and annotation. The reads were annotated with the Genbank database for taxonomy and the Subsystem database for functional genes. Annotations were filtered with a minimum identity of 80%, a minimum aligned length of 20 bp and an E-value cutoff of 1×10^{-5}, which is stricter than the default value. In the output files of MG-RAST, the annotations of functional genes were listed as enzymes encoded by the genes and classified to three levels of broad categories (level 1, level 2 and level 3) in the Subsystem database.

2.3. Statistical analyses

Variations in the relative abundances of functional genes and phylogenetic groups among the samples were analyzed with both univariate and multivariate methods and visualized with the R program. The sequences were rarefied separately for each functional category (level 1, level 2 and function) to the minimum number of sequences among all samples with the 'rrarefy' function in R package 'vegan' (Jari Oksanen et al, 2015). The sequences for phylum and species were rarefied to the minimum number of taxonomic annotations across all samples. The Bray-Curtis distances between sample pairs were calculated based on species composition and visualized with non-metric multidimensional scaling (NMDS) using the 'metaMDS' function in package 'vegan' (Jari Oksanen et al., 2015). The phyla significantly correlated with the ordinations were selected with function 'envfit' and plotted in NMDS figures, with the R^2 and false discovery rate (FDR) of correlations listed in Fig. 1c. The FDR values were calculated according to Benjamini and Hochberg (1995). The Bray-Curtis distances were also calculated from functional gene abundances and visualized with NMDS, and the level 1 categories that significantly correlated with the ordinations were included in Fig. 1b. The relative abundances of functional categories and taxa were averaged for each site, z-transformed and visualized as heatmaps using the R package 'pheatmap'.

To analyze the variations of functional genes and taxa across chronosequence ages, linear models were applied to test the significance of relationships between chronosequence ages and taxa/function with R function 'lm'. In the analysis of the differential abundance of genera, species and functional genes involved in N cycle and greenhouse gas emission across chronosequence age, the sequence count was scaled as the relative abundance multiplied by the minimum total sequence count among all samples. This was used rather than rarefaction to correct for the different sequencing depths because less abundant groups are more sensitive to data removal in rarefaction (McMurdie and Holmes, 2014; Weiss et al., 2017). The linkages between taxonomic groups, functional genes and previously observed



c

Taxa	Abbreviations	R^2	FDR
Acidobacteria	Acido	0.66	0.005**
Actinobacteria	Actino	0.81	0.004**
Bacteroidetes	Bacter	0.59	0.007**
Verrucomicrobia	Ver	0.65	0.007**
Alphaproteobacteria	AlphaP	0.52	0.016*
Betaproteobacteria	BetaP	0.22	0.24
Deltaproteobacteria	DeltaP	0.62	0.007**
Gammaproteobacteria	GammaP	0.76	0.004**

d

Level1 functional category	Abbreviation	R^2	FDR
Amino Acids and Derivatives	AAD	0.47	0.034 *
Carbohydrates	CH	0.31	0.13
Cell Division and Cell Cycle	CDCC	0.48	0.020 *
Cell Wall and Capsule	CWC	0.66	0.002 **
Cofactors, Vitamins, Prosthetic Groups, Pigments	CVPP	0.88	0.002 **
DNA Metabolism	DM	0.21	0.27
Fatty Acids, Lipids, and Isoprenoids	FLI	0.68	0.006 **
Membrane Transport	MT	0.15	0.39
Metabolism of Aromatic Compounds	MAC	0.15	0.39
Nitrogen Metabolism	NM	0.85	0.002 **
Nucleosides and Nucleotides	NN	0.57	0.014 *
Phages, Prophages, Transposable elements, Plasmids	PPTP	0.80	0.002 **
Protein Metabolism	PM	0.70	0.002 **
RNA Metabolism	RM	0.22	0.28
Respiration	Res	0.77	0.002 **
Stress Response	SR	0.83	0.002 **
Virulence, Disease and Defense	VDD	0.85	0.002 **

Fig. 1. The microbial community shift was visualized with NMDS based on (a) species composition and (b) functional gene composition. (c) Major phyla (relative abundance > 3%) were fit to the ordinations based on species composition and those with significant correlations were plotted in (a) (false discovery rate (FDR) < 0.05). (d) Major level 1 categories (relative abundance > 1.5%) were fit to the ordinations based on functional gene composition and those with significant correlations were plotted in (b) (FDR < 0.05). (chronosequence ages 6, 12, 22 and 31: years since reforestation at sampling time; UM: nearby unmined sites as control).

ecosystem function parameters (Avera et al., 2015) involved in C and N cycle were estimated with linear models using the R function 'lm'. The analyzed ecosystem function parameters include total microbial biomass (TMB), substrate induced respiration (SIR), NH_4^+ and NO_3^- concentrations in soil, and emission rates of CO_2, CH_4 and N_2O. Within the ecosystem data collected by Avera et al. (2015), only data collected in September was used for studying the linkages to ensure both datasets were collected in the same month and minimize the impact of seasonal variations. All P values were adjusted to FDR (Benjamini and Hochberg, 1995) in order to minimize Type I errors when testing multiple hypotheses.

2.4. Accession numbers

The sequences can be accessed with project ID mgp16379 in the MG-RAST database.

3. Results

3.1. Shotgun metagenome sequencing results

Shotgun metagenome sequencing resulted in approximately 227 million total sequences, averaging about 15 million per sample. Among all the sequences, 2.81% sequences failed to pass quality control, 2.71% were identified as neither rRNA nor proteins, 6.61% were identified as rRNA, 20.8% were identified as annotated proteins, and 67.1% were

identified as unknown proteins. The percentages are similar to those reported in other soil metagenome studies (Fierer et al., 2012a, 2012b; Leff et al., 2015).

3.2. Taxonomic and functional gene abundances

To understand microbial changes with regard to both taxa and functions, we calculated the Bray-Curtis distances between sample pairs based upon both species and functional gene composition, respectively, and visualized the patterns with NMDS. As shown in Fig. 1a, the patterns of β-diversity based upon species composition generally transition along the chronosequence ages from 6 yr to 31 yr, which is consistent with previously published patterns of these same communities characterized with 16S rRNA gene amplicon sequencing (Sun et al., 2017). The relative abundances of the phyla Actinobacteria, Acidobacteria, Bacteroidetes, Verrucomicrobia, Alphaproteobacteria, Deltaproteobacteria and Gammaproteobacteria were significantly correlated with the ordinations of samples (FDR = 0.004–0.016, R^2 = 0.52–0.81) (Fig. 1a and c), with Actinobacteria and Bacteroidetes being more abundant in earlier ages and Acidobacteria more abundant in later ages. There were also relationships with ecosystem function parameters measured by Avera et al. (2015). Specifically, TMB, SIR, NH_4^+ and emission rates of CH_4 and N_2O were all significantly correlated with the microbial community structure (FDR = 0.0053–0.017, R^2 = 0.45–0.76) (Fig. S1a). TMB and SIR were generally higher in late stages. In contrast, soil NH_4^+ concentrations and emission rates of CH_4 and N_2O were higher in early stages, although the CH_4 relationship was largely driven by its highly negative emission (consumption) in the unmined sites.

In contrast, the ordination plot based on composition of functional genes shows a different pattern (Fig. 1b); the unmined sites were more similar to 22 yr sites and shifted away from the overall 6–31 yr trajectory. The changes in relative abundance of some level 1 function categories were significantly correlated with the ordinations (FDR = 0.002–0.034, R^2 = 0.47–0.88), including Amino Acids and Derivatives (AAD); Cell Division and Cell Cycle (CDCC); Cell Wall and Capsule (CWC); Cofactors, Vitamins, Prosthetic Groups, Pigments (CVPP); Fatty Acids, Lipids, and Isoprenoids (FLI); Nitrogen Metabolism (NM); Nucleosides and Nucleotides (NN); Phages, Prophages, Transposable elements, Plasmids (PPTP); Protein Metabolism (PM); Respiration (Res); Stress Response (SR); and Virulence, Disease and Defense (VDD) (Fig. 1b and d). PPTP, SR and VDD increased with chronosequence ages, while CVPP, CWC and FLI decreased. Res was more abundant at unmined sites, while AAD, CDCC, NM, NN and PM were less abundant. With regards to ecosystem function, correlations of TMB, SIR and CH_4 fluxes with functional gene profiles were marginally significant (FDR = 0.049–0.054, R^2 = 0.47–0.59), with higher TMB, SIR and CH_4 consumption associated with unmined sites (Fig. S1b).

3.3. Patterns of high level taxonomic change

Based on similarities to entries in the MG-RAST database, 97.64% of sequences were classified as Bacteria, 0.59% were classified as Archaea, 1.61% were classified as Eukaryota and 0.02% were classified as viruses. The dominant microbial phyla or proteobacterial classes were Alphaproteobacteria, Actinobacteria and Betaproteobacteria, which respectively represent 34.1%, 18.5% and 12.5% of all taxonomic annotations (Fig. 2). Acidobacteria was less abundant (6.77%) in the shotgun metagenome database than it was among amplified 16S rRNA sequences (21.5%) (Sun et al., 2017), while Alphaproteobacteria were abundant when characterized with both methods.

Multiple taxa with relative abundances > 0.5% varied across chronosequence ages at phylum and proteobacterial class levels (Fig. 2). Actinobacteria were more abundant at 5 yr sites, while Alphaproteobacteria were more abundant at the unmined reference sites. More generally, the relative abundance of Bacteroidetes was higher at earlier ages while that of Acidobacteria was higher at later ages. To further test these trends, linear regression was performed to quantify the significance of the taxonomic variations across ages (6–31 yr) (Fig. 2 and Fig. S2). The relative abundances of Actinobacteria decreased with age from 27.1% at 6 yr to 13.2% in 31 yr (FDR = 0.032, R^2 = 0.59), while the relative abundances of Acidobacteria increased from 4.0% at 6 yr to 11.2% at 31 yr (FDR = 0.047, R^2 = 0.52) (Fig. 2). Among less abundant phyla, Euryarchaeota, Aquificae, Nitrospirae, Thermotogae, Deferribacteres and Zetaproteobacteria all increased with age (FDR = 0.015–0.047, R^2 = 0.52–0.72), with the strongest effect sizes observed for Zetaproteobacteria and Nitrospirae (Fig. S2). Within the dominant Alphaproteobacteria, order-level variation in this class was also analyzed (Fig. S3). The relative abundances of Caulobacterales and Sphingomonadales decreased with age (FDR = 0.0027–0.047, R^2 = 0.43–0.72), while that of Rhodospirillales increased significantly (FDR = 0.0027, R^2 = 0.69).

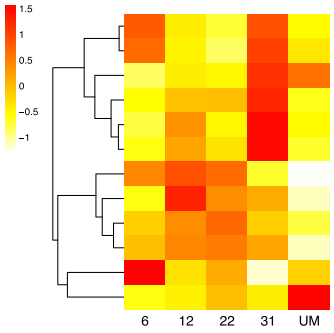
3.4. Patterns of broad functional change

Functional genes were clustered into 21 level 1 categories based on the Subsystem database. The dominant categories were Clustering-based subsystems (CBS); Carbohydrates (CH); Protein Metabolism (PM); and Amino Acids and Derivatives (AAD) – representing 14.2%, 11.8%, 10.4% and 10.3% of the total annotated genes, respectively (Fig. 3). Among all categories, Virulence, Disease and Defense (VDD); Stress Response (SR); Phages, Prophages, Transposable elements, and Plasmids (PPTP); and Secondary Metabolism (SM) were more abundant in later stages (Fig. 3). Photosynthesis genes had high relative abundances at 6 yr sites, while the relative abundances of genes encoding for Potassium metabolism and Metabolism of Aromatic Compounds (MAC) were higher at unmined sites. Among these categories, the relative abundances of PPTP, SR, VDD, SM and Membrane Transport (MT) increased with age (FDR = 0.035–0.077, R^2 = 0.44–0.62), while the relative abundances of FLI and CVPP decreased (FDR = 0.012–0.040, R^2 = 0.53–0.73) (Fig. 3 and Fig. S4). Changes in FLA and SR had the highest effect sizes out of significant categories. It is important to note, however, that the significant slopes observed for some categories (e.g., CVPP, FLI, and MT) within the chronosequence plots did not hold true in the unmined reference plots (Fig. 3).

The analysis of linkages between functional categories and ecosystem parameters revealed that Nucleosides and Nucleotides (NN) is significantly related with TMB (FDR = 0.033, R^2 = 0.58) and VDD is significantly related with SIR (FDR = 0.033, R^2 = 0.57) (Table S1). Among level 2 categories, genes that significantly decreased include: Alanine, serine, and glycine; Aminosugars; Detoxification; Monosaccharides; NAD and NADP; Phospholipids; and RNA processing and modification. Categories that significantly increased include: Fermentation; Oxidative stress; Polysaccharides; Phages, Prophages; Programmed cell death and toxin-antitoxin systems; Cold shock; Pathogenicity islands; Resistance to antibiotics and toxic compounds; and Translation. (FDR = 0.020–0.058, R^2 = 0.47–0.66) (Fig. S5).

3.5. Changes related to nitrogen cycling

To investigate potential changes in microbial nitrogen cycling with reforestation, we conducted more targeted analyses focused on the relative abundances of taxa and functional genes known to affect nitrogen transformations. For nitrification, there were increases in ammonia-oxidizing bacteria (AOB) from 0.61% to 0.97% in 6–31 yr plots and nitrite-oxidizing bacteria (NOB) from 2.5% to 3.3% from 6 yr to 31 yr plots (FDR = 0.050 and 0.047, R^2 = 0.41 and 0.47). However, the relative abundance of NOB increased further to 4.2% in the unmined sites, while AOB was the highest at 31-yr sites and decreased to 0.63% in unmined sites (Fig. 4 and Fig. S6). There was additional complexity within functional groups, with specific taxa trending differently with chronosequence age. For example, among ammonia oxidizers, *Nitrosomonas, Nitrosospira* and *Nitrosococcus* (AOB) increased significantly



Phylum	Mean Abundance (Relative %)	R^2	Slope (± SE)	FDR
Cyanobacteria	1.53E4 (0.97%)	0.0014	8.41 ± 7.05E1	0.95
Chloroflexi	2.22E4 (1.41%)	0.013	5.22E1 ± 1.45E2	0.88
Acidobacteria	1.07E5 (6.77%)	0.52	3.85E3 ± 1.18E3	0.047 *
Firmicutes	3.54E4 (2.25%)	0.37	2.99E2 ± 1.24E2	0.11
Verrucomicrobia	5.52E4 (3.51%)	0.37	8.82E2 ± 3.63E2	0.11
Deltaproteobacteria	6.89E4 (4.37%)	0.40	9.68E2 ± 3.78E2	0.11
Bacteroidetes	5.88E4 (3.73%)	0.31	-8.28E2 ± 3.90E2	0.14
Betaproteobacteria	1.97E5 (12.54%)	0.0093	3.45E2 ± 1.13E3	0.88
Planctomycetes	3.81E4 (2.42%)	0.00010	4.13 ± 1.29E2	0.98
Gammaproteobacteria	9.42E4 (5.98%)	0.0064	7.56E1 ± 2.98E2	0.88
Actinobacteria	2.91E5 (18.47%)	0.59	-6.74E3 ± 1.79E3	0.032 *
Alphaproteobacteria	5.37E5 (34.11%)	0.025	6.64E2 ± 1.31E3	0.78

Fig. 2. The variations of phyla or classes (within Proteobacteria) over chronosequence ages (relative abundance > 0.5%) (chronosequence ages 6, 12, 22 and 31: years since reforestation when sampled; UM: nearby unmined sites as control). The heatmap shows the age variations. The key shows the z-scores of the relative abundances. The relative abundance, variance explained (R^2), regression slope and false discovery rate (FDR) of the linear regression with chronosequence age were shown in the table (• indicates FDR < 0.1, * indicates FDR < 0.05 and ** indicates FDR < 0.01).

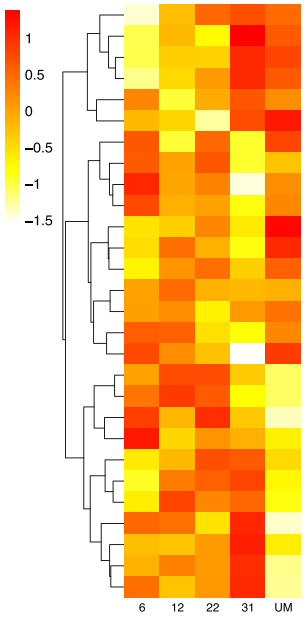
from 6 to 31 yr sites (FDR = 0.036–0.076, R^2 = 0.32–0.53), with similar relative abundances and effect sizes, but they were all less abundant at unmined sites (Fig. 4a and Fig. S6). In contrast, *Nitrosopumilus* (the most abundant ammonia-oxidizing archaea (AOA)) did not change significantly from 6 to 31 yr but increased from 0.0074% at previously mined sites to 0.020% in unmined sites (Fig. 4a).

Among nitrite oxidizers, the relative abundance of *Nitrobacter* increased significantly from 2.5% to 3.2% at 6–31 yr sites (FDR = 0.050, R^2 = 0.43) and increased further to 4.1% in unmined sites. Meanwhile *Nitrococcus* also increased significantly from 0.069% to 0.11% at 6–31 yr sites (FDR = 0.0097, R^2 = 0.68) but then had a mean abundance of only 0.095% at unmined sites (Fig. 4a and Fig. S6). Overall, *Nitrobacter* was the dominant NOB, typically an order of magnitude more abundant than others, so its abundance mainly drove the total NOB pattern. Among functional genes, nitrate reductase (EC 1.7.99.4), which is also known as nitrite oxidoreductase and catalyzes both nitrification and denitrification, was more abundant at the earlier stages (Fig. 4b).

The high phylogenetic diversity of denitrifiers, with > 50 genera suggested to include this group (Shapleigh, 2013; Wang et al., 2014) and the broad distribution of nitrite reductase genes make it difficult to provide a complete list of organisms involved in denitrification. To overcome this challenge, genes involved in denitrification, rather than taxa, are typically targeted to estimate the potential activity of denitrifiers in the environment. Across the chronosequence ages, changes in sequence abundances suggest more active denitrification in earlier stages. For example, genes encoding the nitrite reductase [NAD(P)H] large subunit (EC 1.7.1.4) and nitrite transporters from formate/nitrite family decreased significantly from 6 to 31 yr sites (FDR = 0.028, R^2 = 0.63 and 0.66) (Fig. 4b and Fig. S7), with transporter genes showing a stronger effect size. With regard to measured ecosystem processes, however, no taxonomic changes were correlated with concentrations of soil NH_4^+ or NO_3^- concentrations, nor N_2O emission rates. With regard to N_2O production, nitric oxide reductase and nitrous oxide reductase did not show significant trends across ages (Fig. 5b). The nitric oxide reductase activation protein NorE and N_2O emission were the only functional gene and N cycle parameter pair that were marginally correlated (FDR = 0.057, R^2 = 0.46) (Tables S3 and S4).

3.6. Changes related to methane cycling

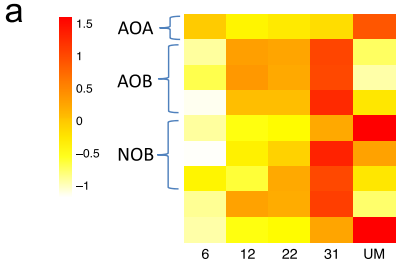
Methanotrophs, which metabolize methane for carbon and energy, comprised 1.1% of all classified sequences at the genus level, while the methanogens, which use CO_2 as an electron acceptor during anaerobic



Level1	Mean Abundance (Relative %)	R^2	Slope (± SE)	FDR
Secondary Metabolism	8.82E3 (0.89%)	0.55	94.3 ± 27.1	0.040*
Virulence, Disease and Defense	2.68E4 (2.69%)	0.55	216.7 ± 62.2	0.040*
Phages, Prophages, Transposable elements, Plasmids	1.56E4 (1.57%)	0.49	76.9 ± 24.9	0.053•
Stress Response	2.98E4 (2.99%)	0.62	119.9 ± 30.0	0.035*
DNA Metabolism	4.20E4 (4.22%)	0.090	36.7 ± 36.8	0.47
Respiration	3.74E4 (3.76%)	0.082	49.7 ± 52.8	0.47
Regulation and Cell signaling	9.98E4 (10.0%)	0.11	-4.4 ± 39.5	0.46
Iron acquisition and metabolism	1.64E3 (0.16%)	0.19	-8.5 ± 5.7	0.35
Cofactors, Vitamins, Prosthetic Groups, Pigments	5.93E4 (5.95%)	0.53	248.2 ± 73.6	0.040*
Miscellaneous	6.06E4 (6.09%)	0.26	-169.4 ± 90.9	0.23
Potassium metabolism	4.95E3 (0.50%)	0.0026	3.1 ± 18.9	0.88
Metabolism of Aromatic Compounds	1.66E4 (1.67%)	0.12	-18.5 ± 15.6	0.46
Sulfur Metabolism	1.45E4 (1.46%)	0.024	10.4 ± 20.8	0.70
Carbohydrates	1.18E5 (11.85%)	0.038	-51.5 ± 81.6	0.66
Motility and Chemotaxis	7.63E3 (0.77%)	0.0091	-5.1 ± 16.7	0.80
Cell Wall and Capsule	1.64E4 (1.65%)	0.32	-150.2 ± 69.7	0.20
Fatty Acids, Lipids, and Isoprenoids	2.68E4 (2.69%)	0.73	-174.9 ± 34.0	0.012*
Amino Acids and Derivatives	1.03E5 (10.35%)	0.031	-23.1 ± 40.7	0.68
Clustering-based subsystems	1.42E5 (14.22%)	0.28	-95.2 ± 47.7	0.23
Nucleosides and Nucleotides	3.38E4 (3.40%)	0.086	-33.2 ± 34.2	0.47
Photosynthesis	1.47E3 (0.15%)	0.13	-5.9 ± 4.8	0.46
Phosphorus Metabolism	1.13E4 (1.13%)	0.23	24.7 ± 14.3	0.27
Membrane Transport	2.70E4 (2.71%)	0.44	49.2 ± 17.7	0.077•
Nitrogen Metabolism	1.77E4 (1.78%)	0.092	39.4 ± 39.5	0.47
Dormancy and Sporulation	1.89E3 (0.19%)	0.013	1.1 ± 3.2	0.78
Protein Metabolism	1.04E5 (10.40%)	0.26	245.7 ± 131.0	0.23
Cell Division and Cell Cycle	1.61E4 (1.61%)	0.18	31.4 ± 21.4	0.35
RNA Metabolism	4.15E4 (4.17%)	0.11	28.5 ± 25.7	0.46

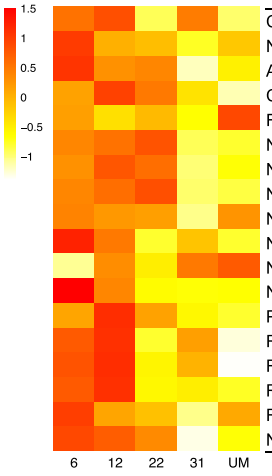
Fig. 3. The variations of level 1 functional categories over chronosequence ages (chronosequence ages 6, 12, 22 and 31: years since reforestation when sampled; UM: nearby unmined sites as control). The heatmap shows the age variations. The key shows the z-scores of the relative abundances. The relative abundance, variance explained (R^2), regression slope and false discovery rate (FDR) of the linear regression with chronosequence age were shown in the table (• indicates FDR < 0.1, * indicates FDR < 0.05 and ** indicates FDR < 0.01).

a



Genus	Mean Abundance (Relative %)	R^2	Slope (± SE)	FDR
Nitrosopumilus	158.2 (0.010%)	0.0067	-0.36 ± 1.4	0.80
Nitrosomonas	3.38E3 (0.21%)	0.39	58.8 ± 23.3	0.050*
Nitrosospira	4.83E3 (0.30%)	0.32	79.9 ± 36.5	0.076•
Nitrosococcus	3.98E3 (0.25%)	0.53	51.2 ± 15.2	0.036*
Nitrobacter	4.81E4 (3.03%)	0.43	413.7 ± 150.1	0.050*
Nitrococcus	1.42E3 (0.089%)	0.68	25.2 ± 5.5	0.0097**
Nitrospira	179.1 (0.011%)	0.29	7.6 ± 3.8	0.090•
AOB	1.22E4 (0.77%)	0.41	190.0 ± 72.8	0.050*
NOB	4.97E4 (3.13%)	0.47	446.5 ± 150.4	0.047*

b



Functional gene	Mean Abundance (Relative %)	R^2	Slope (± SE)	FDR
Copper-containing nitrite reductase (EC 1.7.2.1)	85.7 (9.09E-03%)	0.060	-1.04 ± 1.30	0.50
NnrU family protein	15.0 (1.59E-03%)	0.35	-1.22 ± 0.52	0.14
Assimilatory nitrate reductase large subunit (EC:1.7.99.4)	916.5 (9.72E-02%)	0.52	-12.14 ± 3.65	0.067•
Cytochrome c-type protein NapC	53.0 (5.62E-03%)	0.12	-0.49 ± 0.43	0.41
Ferredoxin--nitrite reductase (EC 1.7.7.1)	210.1(2.22E-02%)	0.067	-1.26 ± 1.49	0.50
Nitrate ABC transporter, ATP-binding protein	211.1 (2.24E-02%)	0.28	-2.87 ± 1.45	0.16
Nitrate ABC transporter, nitrate-binding protein	491.2 (5.21E-02%)	0.27	-6.84 ± 3.55	0.16
Nitrate ABC transporter, permease protein	198.7 (2.11E-02%)	0.31	-2.69 ± 1.27	0.16
Nitrate reductase cytochrome c550-type subunit	17.7 (1.88E-03%)	0.30	-0.34 ± 0.17	0.16
Nitrate/nitrite transporter	171.5 (1.82E-02%)	0.35	-13.44 ± 5.84	0.14
Nitrite reductase probable electron transfer 4Fe-S subunit (EC 1.7.1.4)	17.4 (1.85E-03%)	0.17	0.51 ± 0.36	0.31
Nitrite transporter from formate/nitrite family	82.3 (8.73E-03%)	0.66	-2.46 ± 0.56	0.028*
Periplasmic nitrate reductase precursor (EC 1.7.99.4)	275.8 (2.93E-02%)	0.19	-3.41 ± 2.20	0.26
Respiratory nitrate reductase alpha chain (EC 1.7.99.4)	205.0 (2.18E-02%)	0.27	-3.90 ± 2.02	0.16
Respiratory nitrate reductase beta chain (EC 1.7.99.4)	102.0 (1.08E-02%)	0.41	-2.20 ± 0.84	0.11
Respiratory nitrate reductase gamma chain (EC 1.7.99.4)	26.5 (2.81E-03%)	0.42	-1.06 ± 0.39	0.11
Response regulator NasT	76.5 (8.11E-03%)	0.43	-1.44 ± 0.53	0.11
Nitrite reductase [NAD(P)H] large subunit (EC 1.7.1.4)	792.1 (8.40E-02%)	0.63	-19.2 ± 4.68	0.028*

Fig. 4. The variations of microbial functional groups, genera and functional genes involved in N cycle over chronosequence ages (chronosequence ages 6, 12, 22 and 31: years since reforestation when sampled; UM: nearby unmined sites as control). (a) AOB, NOB and major genera of AOB and NOB over chronosequence ages (relative abundance > 0.005%). (b) Major functional genes involved in N cycle over chronosequence ages (relative abundance > 0.001%). The heatmap shows the age variations. The key shows the z-scores of the relative abundances. The relative abundance, variance explained (R^2), regression slope and false discovery rate (FDR) of the linear regression with chronosequence age were shown in the table (• indicates FDR < 0.1, * indicates FDR < 0.05 and ** indicates FDR < 0.01).

respiration and produce methane, comprised only 0.12% of sequences. All characterized methanotrophs increased significantly by approximately two-to three-fold with chronosequence age (FDR = 0.0013–0.016, R^2 = 0.49–0.78) (Fig. 5a and Fig. S8). The highest relative abundances of *Methylocella*, *Methylocystis* and *Methylosinus* were found at the unmined sites, as 0.71%, 0.27% and 0.28%, respectively. In contrast, *Methylobacter* and *Methylococcus* abundances were highest at 31 yr sites at 0.095% and 0.20%, respectively (Fig. 5a). However, *Methylocella*, *Methylocystis* and *Methylosinus* comprised 76.7% of all methanotrophs, so they mainly drove the overall methanotroph trend. For methanogens, *Methanothermobacter, Methanoculieus, Methanospirillum, Methanoregula* and *Methanosarcina* increased significantly with similar effect sizes along the chronosequence age (FDR = 0.0024–0.018, R^2 = 0.47–0.70) (Fig. S8), but their abundances were much lower than the methanotrophs. The relative abundance changes of *Methylocella*, *Methylocystis* and *Methylosinus* were all significantly negatively related with the emission rates of CH_4 (FDR = 0.00002–0.0002, R^2 = 0.76–0.85) (Table S6).

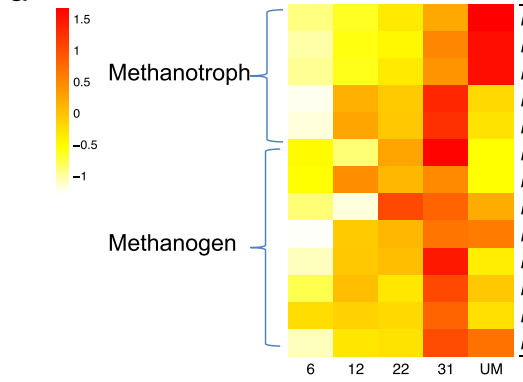
Among functional genes involved in methane cycling, no methyl-coenzyme M reductase genes (the primary gene involved in methanogenesis) were detected, which suggests that net fluxes of methane depended primarily on changes in methanotrophs. Genes related to methane monoxygenase, the functional gene metabolizing methane during methanotrophy, did not change significantly along the chronosequence age. However, combined across all four categories, they did increase sharply from 0.013% in 31 yr sites to 0.029% in unmined sites (Fig. 5b), which further supports that methane consumption was only active at the unmined sites. Moreover, the changes in relative abundances of methane monoxygenase genes were significantly associated with CH_4 emission rates (FDR = 0.0012–0.0036, R^2 = 0.49–0.65) (Table S7).

4. Discussion

4.1. Taxonomic changes

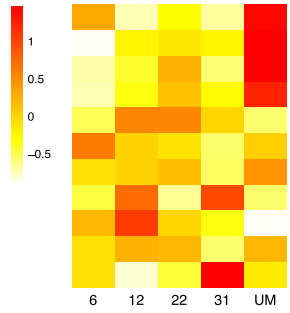
The β-diversity patterns and taxonomic changes among these communities, as characterized by shotgun metagenome sequencing, were generally similar to those previously identified using 16S rRNA gene amplicon sequencing (Sun et al., 2017). Based on Bray-Curtis distances among the samples, similarity of microbial communities was again related to chronosequence age, with the community structure becoming more similar to unmined reference soils as age increased (Fig. 1a). The shotgun metagenome approach, which presumably provides more accurate data on relative abundances without potential amplification biases, verified the previously published conclusions. Strongest changes were observed as nearly doubled Acidobacteria and Nitrospirae (Noll et al., 2005; Davis et al., 2011; Hui et al., 2017) and nearly halved abundances of Actinobacteria (Fierer et al., 2012a; Ramirez et al., 2012). These changes support the shift from copiotrophic to oligotrophic groups described previously (Sun et al., 2017). Overall, chronosequence age explained more than 50% of the variation (R^2 > 0.52) in these taxa, suggesting strong associations between ecosystem development and the copiotrophy-oligotrophy transition. However, it is important to note that the temporal trends observed for

a



Genus	Mean Abundance (Relative %)	R^2	Slope (± SE)	FDR
Methylocella	7.13E3 (0.45%)	0.76	107.3 ± 18.9	0.00131**
Methylocystis	2.62E3 (0.17%)	0.78	57.6 ± 9.7	0.00131**
Methylosinus	3.05E3 (0.19%)	0.73	43.9 ± 8.5	0.00176**
Methylobacter	1.21E3 (0.076%)	0.56	18.5 ± 5.2	0.00963**
Methylococcus	2.68E3 (0.17%)	0.49	32.8 ± 10.5	0.0159*
Methanothermobacter	1.45E2 (0.0091%)	0.70	3.91 ± 0.82	0.00242**
Methanococcus	1.09E2 (0.0069%)	0.077	0.60 ± 0.66	0.381
Methanoculleus	96.7 (0.0061%)	0.62	1.88 ± 0.46	0.00486**
Methanospirillum	1.31E2 (0.0082%)	0.47	1.77 ± 0.59	0.0179*
Methanoregula	2.00E2 (0.013%)	0.64	3.86 ± 0.91	0.00449**
Methanosphaerula	1.26E2 (0.0080%)	0.26	1.85 ± 0.99	0.108
Methanococcoides	1.16E2 (0.0073%)	0.12	1.01 ± 0.85	0.285
Methanosarcina	9.25E2 (0.058%)	0.54	14.6 ± 4.3	0.0111*

b



Functional gene	Mean Abundance (Relative %)	R^2	Slope (± SE)	FDR
Methane monooxygenase component A alpha chain (EC 1.14.13.25)	133.1 (1.41E-02%)	0.12	-1.13 ± 0.96	0.57
Methane monooxygenase component A beta chain (EC 1.14.13.25)	29.3 (3.11E-03%)	0.11	0.48 ± 0.43	0.57
Methane monooxygenase component C (EC 1.14.13.25)	26.5 (2.81E-03%)	0.029	0.23 ± 0.42	0.65
Methane monooxygenase regulatory protein B	12.4 (1.32E-03%)	0.086	0.15 ± 0.16	0.59
Nitric oxide -responding transcriptional regulator NnrR (Crp/Fnr family)	15.9 (1.69E-03%)	0.022	0.11 ± 0.23	0.65
Nitric oxide reductase activation protein NorD	16.8 (1.78E-03%)	0.16	-0.47 ± 0.34	0.57
Nitric oxide reductase activation protein NorQ	45.9 (4.87E-03%)	0.021	-0.14 ± 0.30	0.65
Nitric-oxide reductase (EC 1.7.99.7), quinol-dependent	120.2 (1.28E-02%)	0.079	2.38 ± 2.57	0.59
Nitric-oxide reductase subunit B (EC 1.7.99.7)	16.9 (1.79E-03%)	0.13	-0.39 ± 0.32	0.57
Nitrous oxide reductase maturation protein NosR	15.3 (1.62E-03%)	0.042	-0.21 ± 0.32	0.65
Nitrous-oxide reductase (EC 1.7.99.6)	14.4 (1.52E-03%)	0.32	1.14 ± 0.53	0.57

Fig. 5. The variations of microbial genera and functional genes involved in greenhouse gas emission over chronosequence ages (chronosequence ages 6, 12, 22 and 31: years since reforestation when sampled; UM: nearby unmined sites as control). (a) Major methanotrophs and methanogens over chronosequence ages (relative abundance > 0.005%). (b) Major functional genes involved in methane and nitrous oxide production over chronosequence ages (relative abundance > 0.001%). The heatmap shows the age variations. The key shows the z-scores of the relative abundances. The relative abundance, variance explained (R^2), regression slope and false discovery rate (FDR) of the linear regression with chronosequence age were shown in the table (• indicates FDR < 0.1, * indicates FDR < 0.05 and ** indicates FDR < 0.01).

several important taxa along the chronosequence plots did not hold true in the unmined reference plots. Thus, there is a need to further study the continued development of these communities beyond the first few decades.

Among the 7 orders of the dominant alphaproteobacteria, Caulobacterales and Sphingomonadales decreased while Rhodospirillales increased, with similarly high levels of variance (43.3–71.8%) explained by chronosequence age. Sphingomonadales are known to interact with plants and utilize root exudates (el Zahar Haichar et al., 2008), so development of the tree canopy and reduced cover of grasses in later ages could explain the decrease of Sphingomonadales. Root biomass has been shown to decrease with reduced grass dominance in other systems (Reich et al., 2001) and grass rooting occurs more closely to the surface in contrast to deeper tree rooting in mine soils (Clark and Zipper, 2016). In contrast, Rhodospirillales are more abundant in forest soils than in grassland soils (Nacke et al., 2011) and are generally oligotrophic and negatively correlated with labile soil nutrients such as dissolved nitrogen and organic carbon (King et al., 2010). Because the rhizosphere is typically nutrient-rich (Philippot et al., 2013), the change to deeper rooting vegetation could also reduce labile nutrients in the top 10 cm of soil sampled in this study. Such a shift would also further explain the increase of oligotrophs in shallow soil depths with recovery time. Overall, our results support that the transition from grass-to tree-dominated vegetation and changes in soil nutrients are important factors driving microbial community changes during forest ecosystem recovery, which has also been observed during secondary succession (Cline and Zak (2015)).

4.2. Carbon cycling

Changes in relative abundances of functional genes also indicated that resource availability and microbial interactions are key to the responses of the soil microbiome during forest succession. With regard to carbon metabolism, photosynthesis-related genes were most abundant at 6 yr sites, where the undeveloped tree canopy allowed the most light to reach the soil surface. Among heterotrophs, functional gene changes also indicate that soil communities were adapting to a shift from labile to recalcitrant C compounds with chronosequence age. For example, there was a relative increase in genes involved in polysaccharide metabolism and a decrease in those involved in monosaccharide metabolism, with 64% and 47% of their variability, respectively, explained by chronosequence ages (Fig. S5). Genes involved in the metabolism of aromatic compounds (MAC) were most abundant at unmined sites (Fig. 3), where the amount of organic carbon was also the highest (Avera et al., 2015). A shift from copiotrophy to oligotrophy is also supported by decreases in genes involved in the level 2 cell division category with chronosequence age (Fig. S5), given that oligotrophs tend to have lower turnover rates. Furthermore, shifts from oligotrophic to copiotrophic groups following N addition can result in decreased respiration rates (Treseder, 2008; Janssens et al., 2010; Liu and Greaver, 2010), which agrees with our observation that chronosequence age was previously associated with increased respiration rates at these sites (Avera et al., 2015). The increased respiration is likely due to more utilization by oligotrophs of the recalcitrant C pool, which is the major fraction of soil organic carbon (Fierer et al., 2012a).

4.3. Biological interactions

Soil is a reservoir of phages, prophages and transposable elements (D'Costa et al., 2007; Luo et al., 2016), and these mobile genetic elements can shape microbial communities (Koskella and Meaden, 2013; Sun and Relman, 2013). For example, sequences annotated as transposable elements increased more than 20% in this chronosequence, which could be a response to stresses on the community such as nutrient deficiency or temperature change (McClintock, 1984; Capy et al., 2000). In addition, transposable elements are essential in the construction and spread of catabolic pathways among bacteria, and allow more rapid adaptation of the bacterial community to new xenobiotics (Top and Springael, 2003; Springael and Top, 2004). Thus, the increase of genes in the PPTP category also likely reflects adaption of the bacterial communities to the appearance of new compounds from changing vegetation inputs.

Abundances of genes involved in Virulence, Disease and Defense (VDD) increased by ~25% in this chronosequence, suggesting that interspecific microbial competition also increased with age (Davies and Davies, 2010; Fierer et al., 2012b). Bacteria can increase production of antibiotics and bacteriocin when detecting nutrient competition (Cytryn, 2013; Cornforth and Foster, 2015). Within the VDD category, the relative abundances of genes categorized as toxins and superantigens, although low, approximately tripled with chronosequence age while resistance to antibiotics and toxic compounds increased around 20%, which agrees with increased competition. With chronosequence ages explaining > 47% of the changes of interaction-related functional categories, increased interactions among microorganisms, both cooperative and competitive, were strongly associated with ecosystem development. This result agrees with the increased complexity of the interaction networks (Sun et al., 2017) and increased rates of microbial activity (Avera et al., 2015) that has been previously reported at these sites.

4.4. Nitrogen cycling

Abundance changes of N cycling organisms indicated shifts in overall N cycling as well as shifts within functional groups. For example, AOB and NOB both increased by > 30% with chronosequence age, suggesting a shift towards increased nitrifiers. However, AOB peaked at 31 yr sites, while AOA were highest at unmined sites. An AOB-AOA shift with increasing age and soil C:N agrees with previous reports that AOB usually inhabit ammonium rich environments, while AOA are less dependent on ammonium concentrations (Ollivier et al., 2011). Although the total N was the highest at unmined sites (Avera et al., 2015), N in soil tends to get embedded in complex organic compounds over time, after which soil nitrogen bioavailability highly depends on the depolymerization of N bound in soil organic matter (SOM) (Finzi and Berthrong, 2005). Similar taxonomic shifts also occurred within the NOB. *Nitrobacter*, *Nitrococcus* and *Nitrospirae* all increased with chronosequence age, but *Nitrobacter* was by far the dominant N oxidizing group, particularly at unmined sites, with a mean abundance more than 10-fold greater than any other group. *Nitrobacter* has a lower nitrite affinity than *Nitrospira*, indicating that *Nitrobacter* is more adapted to environments with higher nitrite concentrations (Blackburne et al., 2007).

In contrast to taxonomic changes, most of the functional genes involved in N cycling did not change significantly with chronosequence age (Fig. 4b). Nitrite oxidoreductase was generally higher in the earlier chronosequence ages but is difficult to interpret because it metabolizes the nitrite-nitrate transformation in both directions. For denitrification, nitrite reduction genes generally decreased with chronosequence age (Fig. 4b), implying higher denitrifying activity in younger ages, which is consistent with previous studies (Kandeler et al., 2006; Brankatschk et al., 2011). However, few of these changes were significant and denitrification can be related to a number of factors including moisture, soil compaction, pH, and available nutrients that are expected to change as reclaimed mine soils develop (Groffman and Tiedje, 1989; Petersen et al., 2008; Čuhel et al., 2010). At these sites, no data on soil compaction were collected and there were no significant associations between nitrite reduction genes and nutrient or moisture measurements.

4.5. Methane cycling

Decreased relative abundances of methanotrophs and methane monooxygenase genes at all sites except the unmined sites agrees with the lack of methane consumption in all mined sites previously observed by Avera et al. (2015). Thus, the lack of methane consumption appears to be due to low recovery of methanotroph populations, rather than changes in activity rates alone. Land use is known to alter methanotrophic communities (MacDonald et al., 1997; Nazaries et al., 2011), and deforested soils can switch to a net source of methane from a net sink (Zerva and Mencuccini, 2005).

Methanotrophs are divided into type-I (e.g. *Methylococcus* and *Methylobacter*) and type-II (e.g. *Methylosinus and Methylocystis*), and we observed a shift from type-I to type-II with age, agreeing with previous observations (Nazaries et al., 2011). Overall, type-II methanotrophs were dominant (77% of methanotrophs) and responded to chronosequence age with larger effect sizes. Thus, they likely drove most of the change in methane monooxygenase gene abundances, given that both reached the highest relative abundances at the unmined sites. Similarly, relative abundances of type-II methanotrophs were significantly correlated with methane oxidation at other reforestation sites (Nazaries et al., 2011).

Although all of the type-II methanotrophs increased significantly with age, their relative abundances were still much higher at unmined sites. Thus, the methanotrophic community was only beginning to recover and will likely take much longer than 31 years to become comparable to undisturbed forest soils. Methanotrophs living at low methane concentrations are extremely vulnerable to disturbances (Priemé et al., 1997), and the recovery time of the soil methane sink after land use change has been previously estimated to be > 100 yr (Smith et al., 2000). Furthermore, the transition of methanotrophs into an inactive state and inhibition from N turnover could further affect the recovery of methane oxidation (Priemé et al., 1997; Ho et al., 2011). Specific factors controlling recovery of soil methane sinks are mostly unknown, but they are likely regulated by complex processes that include not only abiotic factors but also interactions with methanogens and other organisms.

4.6. Linkages between metagenome and ecosystem processes

Our results suggest that the strength of relationships between ecosystem processes and changes in the relevant portion of the metagenome depends on the type of processes examined. For the N cycle, only one marginally significant relationship was revealed between relevant taxa or functional genes and measurements of soil NH_4^+, NO_3^- or N_2O flux. In contrast, methanotrophic taxa, methane monooxygenase genes, and CH_4 fluxes were strongly related. For example, each of the three type-II methanotrophs in Fig. 5 could explain > 75% of the variance in CH_4 emission rates, while methane monooxygenase genes explained > 49% (Tables S6 and S7). In contrast, < 40% of the variance in NH_4^+, NO_3^- concentrations or N_2O flux could be explained by any of the related taxa (Fig. 4; Table S3). Functionally, 45.7% of the variance of N_2O emission rates could be explained by NorE, but explained variance was < 12% for other denitrification genes (Table S4).

Ecosystem process rates are often difficult to associate with changes in microbial community structure or functional gene abundance (Rocca et al., 2015; Hall et al., 2018), which is likely to be compounded for more complex cycles. For example, the N cycle, and particularly denitrification, consists of more biogeochemical transformations and

pathways and is more broadly distributed across phylogenetic groups than the CH_4 cycle. Furthermore, in this study, the magnitude of change or "signal-to-noise ratio" in CH_4 emission rates observed at these sites was much higher than for changes in the N cycle (Avera et al., 2015). Considering all these factors, it is not surprising that associations between metagenomes and ecosystem processes are harder to detect in the N cycle than in the CH_4 cycle.

As sequencing capacity continues to increase, improved associations between metagenomes and ecosystem processes may result from increased sample sizes. However, this may not be enough; Hall et al. (2018) outlined the challenges in linking microbiome changes to ecosystem function directly, particularly in complex cycles and without accounting for intermediate mechanistic steps. A more comprehensive experimental framework focused on individual processes, as they proposed, would provide better understanding of the linkages in complex processes. Based on our results, the ease of identifying relationships between the microbiome and biogeochemical cycles is likely related to the number of individual transformations that results in a net flux, the phylogenetic diversity of microbes involved in the process, and the magnitude of changes observed in process rates in a specific environment. It is also important to note that metagenomic approaches are biased towards bacteria because of their dominance in the soil DNA pool. Therefore, other important mediators of soil biogeochemical processes such as fungi, soil animals and plant root biomass are not as well incorporated. Future studies that increase sample size, better integrate soil eukaryotes, and more precisely target specific biogeochemical processes are essential for a better understanding of the relationships between soil communities, genes and processes.

Overall, this work describes both taxonomic and functional changes in soil metagenomes following ecosystem disturbance and elucidates possible mechanisms driving microbial recovery after disturbance and their potential ecological implications. The shifts of microbial taxa and functional genes both indicated that a transition from labile to recalcitrant carbon was important in driving microbial metabolic change. The resource transition likely led to the shift from copiotrophic to oligotrophic groups and increased competition between microorganisms. Furthermore, microorganisms displayed successional patterns not only at the community level, but also within functional groups, such as the transition between AOB and AOA within ammonia oxidizers and the transitions among *Nitrobacter, Nitrospira* and *Nitrococcus* within nitrite oxidizers. Abundances of both taxa and functional genes indicated that the extremely long recovery time of methanotroph populations is the reason that the forest soils did not recover as methane sinks after 31 years. The different relationships revealed between metagenome and ecosystem parameters in two key biogeochemical cycles suggest that the complexity of those cycles and magnitude of observed changes in process rates are limiting factors in identifying broad relationships between complete metagenome datasets and ecosystem function.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

This study was funded by the Powell River Project and the Virginia Agricultural Experiment Station.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.soilbio.2019.05.004.

References

Chapuis-Lardy, L., Wrage, N., Metay, A., Chotte, J.L., Bernoux, M., 2007. Soils, a sink for N2O? A review. Global Change Biology 13, 1–17.
Akselsson, C., Westling, O., 2005. Regionalized nitrogen budgets in forest soils for different deposition and forestry scenarios in Sweden. Global Ecology and Biogeography 14, 85–95.
Avera, B.N., Strahm, B.D., Burger, J.A., Zipper, C.E., 2015. Development of ecosystem structure and function on reforested surface-mined lands in the Central Appalachian Coal Basin of the United States. New Forests 46, 683–702.
Baldrian, P., Kolařík, M., Štursová, M., Kopecký, J., Valášková, V., Větrovský, T., Žifčáková, L., Šnajdr, J., Rídl, J., Vlček, Č., 2012. Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. The ISME Journal 6, 248–258.
Banning, N.C., Gleeson, D.B., Grigg, A.H., Grant, C.D., Andersen, G.L., Brodie, E.L., Murphy, D.V., 2011. Soil microbial community successional patterns during forest ecosystem restoration. Applied and Environmental Microbiology 77, 6158–6164.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 289–300.
Blackburne, R., Vadivelu, V.M., Yuan, Z., Keller, J., 2007. Kinetic characterisation of an enriched Nitrospira culture with comparison to Nitrobacter. Water Research 41, 3033–3042.
Bonan, G.B., 2008. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. Science 320, 1444–1449.
Böttcher, H., Verkerk, P.J., Gusti, M., Havlík, P., Grassi, G., 2012. Projection of the future EU forest CO2 sink as affected by recent bioenergy policies using two advanced forest management models. Gcb Bioenergy 4, 773–783.
Brankatschk, R., Töwe, S., Kleineidam, K., Schloter, M., Zeyer, J., 2011. Abundances and potential activities of nitrogen cycling microbial communities along a chronosequence of a glacier forefield. The ISME Journal 5, 1025–1037.
Bukata, A.R., Kyser, T.K., 2005. Response of the nitrogen isotopic composition of tree-rings following tree-clearing and land-use change. Environmental Science & Technology 39, 7777–7783.
Capy, P., Gasperi, G., Biémont, C., Bazin, C., 2000. Stress and transposable elements: co-evolution or useful parasites? Heredity 85, 101–106.
Clark, E.V., Zipper, C.E., 2016. Vegetation influences near-surface hydrological characteristics on a surface coal mine in eastern USA. Catena 139, 241–249.
Cline, L.C., Zak, D.R., 2015. Soil microbial communities are shaped by plant-driven changes in resource availability during secondary succession. Ecology 96, 3374–3385.
Cornforth, D.M., Foster, K.R., 2015. Antibiotics and the art of bacterial war. Proceedings of the National Academy of Sciences 112, 10827–10828.
Čuhel, J., Šimek, M., Laughlin, R.J., Bru, D., Chèneby, D., Watson, C.J., Philippot, L., 2010. Insights into the effect of soil pH on N2O and N2 emissions and denitrifier community size and activity. Applied and Environmental Microbiology 76, 1870–1878.
Cytryn, E., 2013. The soil resistome: the anthropogenic, the native, and the unknown. Soil Biology and Biochemistry 63, 18–23.
Davies, J., Davies, D., 2010. Origins and evolution of antibiotic resistance. Microbiology and Molecular Biology Reviews 74, 417–433.
Davis, K.E., Sangwan, P., Janssen, P.H., 2011. Acidobacteria, Rubrobacteridae and Chloroflexi are abundant among very slow-growing and mini-colony-forming soil bacteria. Environmental Microbiology 13, 798–805.
De Vries, F.T., Van Groenigen, J.W., Hoffland, E., Bloem, J., 2011. Nitrogen losses from two grassland soils with different fungal biomass. Soil Biology and Biochemistry 43, 997–1005.
Dini-Andreote, F., de Cassia Pereira e Silva, M., Triado-Margarit, X., Casamayor, E.O., van Elsas, J.D., Salles, J.F., 2014. Dynamics of bacterial community succession in a salt marsh chronosequence: evidences for temporal niche partitioning. The ISME Journal 8, 1989–2001.
D'Costa, V.M., Griffiths, E., Wright, G.D., 2007. Expanding the soil antibiotic resistome: exploring environmental diversity. Current Opinion in Microbiology 10, 481–489.
el Zahar Haichar, F., Marol, C., Berge, O., Rangel-Castro, J.I., Prosser, J.I., Balesdent, J., Heulin, T., Achouak, W., 2008. Plant host habitat and root exudates shape soil bacterial community structure. The ISME Journal 2, 1221–1230.
Fearnside, P.M., 2000. Global warming and tropical land-use change: greenhouse gas emissions from biomass burning, decomposition and soils in forest conversion, shifting cultivation and secondary vegetation. Climatic Change 46, 115–158.
Fierer, N., Lauber, C.L., Ramirez, K.S., Zaneveld, J., Bradford, M.A., Knight, R., 2012a. Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. The ISME Journal 6, 1007–1017.
Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S., Gilbert, J.A., Wall, D.H., Caporaso, J.G., 2012b. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proceedings of the National Academy of Sciences 109, 21390–21395.
Fierer, N., Ladau, J., Clemente, J.C., Leff, J.W., Owens, S.M., Pollard, K.S., Knight, R., Gilbert, J.A., McCulley, R.L., 2013. Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. Science 342, 621–624.
Finzi, A.C., Berthrong, S.T., 2005. The uptake of amino acids by microbes and trees in three cold-temperate forests. Ecology 86, 3345–3353.
Fullerton, H., Hager, K.W., McAllister, S.M., Moyer, C.L., 2017. Hidden diversity revealed by genome-resolved metagenomics of iron-oxidizing microbial mats from Lō'ihi Seamount, Hawai'i. The ISME Journal 11, 1900.
Groffman, P.M., Tiedje, J.M., 1989. Denitrification in north temperate forest soils: spatial and temporal patterns at the landscape and seasonal scales. Soil Biology and Biochemistry 21, 613–620.
Hall, E.K., Bernhardt, E.S., Bier, R.L., Bradford, M.A., Boot, C.M., Cotner, J.B., del Giorgio,

P.A., Evans, S.E., Graham, E.B., Jones, S.E., 2018. Understanding how microbiomes influence the systems they inhabit. Nature Microbiology 3, 977.
Ho, A., Lüke, C., Frenzel, P., 2011. Recovery of methanotrophs from disturbance: population dynamics, evenness and functioning. The ISME Journal 5, 750–758.
Hui, N., Jumpponen, A., Francini, G., Kotze, D.J., Liu, X., Romantschuk, M., Strömmer, R., Setälä, H., 2017. Soil microbial communities are shaped by vegetation type and park age in cities under cold climate. Environmental Microbiology 19 (3), 1281–1295.
Janssens, I., Dieleman, W., Luyssaert, S., Subke, J.-A., Reichstein, M., Ceulemans, R., Ciais, P., Dolman, A.J., Grace, J., Matteucci, G., 2010. Reduction of forest soil respiration in response to nitrogen deposition. Nature Geoscience 3, 315–322.
Jari Oksanen, F.G.B., Kindt, Roeland, Legendre, Pierre, Minchin, Peter R., O'Hara, R.B., Simpson, Gavin L., Peter, Solymos, Stevens, M. Henry H., Wagner, Helene, 2015. vegan: community Ecology Package. R package version 2.2-1. http://CRAN.R-project.org/package=vegan.
Kandeler, E., Deiglmayr, K., Tscherko, D., Bru, D., Philippot, L., 2006. Abundance of narG, nirS, nirK, and nosZ genes of denitrifying bacteria during primary successions of a glacier foreland. Applied and Environmental Microbiology 72, 5957–5962.
King, A.J., Freeman, K.R., McCormick, K.F., Lynch, R.C., Lozupone, C., Knight, R., Schmidt, S.K., 2010. Biogeography and habitat modelling of high-alpine bacteria. Nature Communications 1, 53.
Knelman, J.E., Schmidt, S.K., Lynch, R.C., Darcy, J.L., Castle, S.C., Cleveland, C.C., Nemergut, D.R., 2014. Nutrient addition dramatically accelerates microbial community succession. PLoS One 9, e102609.
Kolb, S., 2009. The quest for atmospheric methane oxidizers in forest soils. Environmental Microbiology Reports 1, 336–346.
Koskella, B., Meaden, S., 2013. Understanding bacteriophage specificity in natural microbial communities. Viruses 5, 806–823.
Kumaresan, D., Cross, A.T., Moreira-Grez, B., Kariman, K., Nevill, P., Stevens, J., Allcock, R.J., O'Donnell, A.G., Dixon, K.W., Whiteley, A.S., 2017. Microbial functional capacity is preserved within engineered soil formulations used in mine site restoration. Scientific Reports 7, 564.
Leff, J.W., Jones, S.E., Prober, S.M., Barberán, A., Borer, E.T., Firn, J.L., Harpole, W.S., Hobbie, S.E., Hofmockel, K.S., Knops, J.M., 2015. Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. Proceedings of the National Academy of Sciences 112, 10967–10972.
Liu, L., Greaver, T.L., 2010. A global perspective on belowground carbon dynamics under nitrogen enrichment. Ecology Letters 13, 819–828.
Luo, W., Xu, Z., Riber, L., Hansen, L.H., Sørensen, S.J., 2016. Diverse gene functions in a soil mobilome. Soil Biology and Biochemistry 101, 175–183.
MacDonald, J., Skiba, U., Sheppard, L., Ball, B., Roberts, J., Smith, K., Fowler, D., 1997. The effect of nitrogen deposition and seasonal variability on methane oxidation and nitrous oxide emission rates in an upland spruce plantation and moorland. Atmospheric Environment 31, 3693–3706.
McClintock, B., 1984. The significance of responses of the genome to challenge. Science 226, 792–801.
McMurdie, P.J., Holmes, S., 2014. Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Computational Biology 10, e1003531.
Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., 2008. The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 9, 386.
Nacke, H., Thürmer, A., Wollherr, A., Will, C., Hodac, L., Herold, N., Schöning, I., Schrumpf, M., Daniel, R., 2011. Pyrosequencing-based assessment of bacterial community structure along different management types in German forest and grassland soils. PLoS One 6, e17000.
Nazaries, L., Tate, K.R., Ross, D.J., Singh, J., Dando, J., Saggar, S., Baggs, E.M., Millard, P., Murrell, J.C., Singh, B.K., 2011. Response of methanotrophic communities to afforestation and reforestation in New Zealand. The ISME Journal 5, 1832–1836.
Noll, M., Matthies, D., Frenzel, P., Derakshani, M., Liesack, W., 2005. Succession of bacterial community structure and diversity in a paddy soil oxygen gradient. Environmental Microbiology 7, 382–395.
Ollivier, J., Töwe, S., Bannert, A., Hai, B., Kastl, E.M., Meyer, A., Su, M.X., Kleineidam, K., Schloter, M., 2011. Nitrogen turnover in soil and global change. FEMS Microbiology Ecology 78, 3–16.
Paver, S.F., Hayek, K.R., Gano, K.A., Fagen, J.R., Brown, C.T., Davis-Richardson, A.G., Crabb, D.B., Rosario-Passapera, R., Giongo, A., Triplett, E.W., 2013. Interactions between specific phytoplankton and bacteria affect lake bacterial community succession. Environmental Microbiology 15, 2489–2504.
Petersen, S.O., Schjønning, P., Thomsen, I.K., Christensen, B.T., 2008. Nitrous oxide evolution from structurally intact soil as influenced by tillage and soil water content. Soil Biology and Biochemistry 40, 967–977.
Philippot, L., Andersson, S.G., Battin, T.J., Prosser, J.I., Schimel, J.P., Whitman, W.B., Hallin, S., 2010. The ecological coherence of high bacterial taxonomic ranks. Nature Reviews Microbiology 8, 523–529.
Philippot, L., Raaijmakers, J.M., Lemanceau, P., Van Der Putten, W.H., 2013. Going back to the roots: the microbial ecology of the rhizosphere. Nature Reviews Microbiology 11, 789.
Priemé, A., Christensen, S., Dobbie, K.E., Smith, K.A., 1997. Slow increase in rate of methane oxidation in soils with time following land use change from arable agriculture to woodland. Soil Biology and Biochemistry 29, 1269–1273.
Ramirez, K.S., Craine, J.M., Fierer, N., 2012. Consistent effects of nitrogen amendments on soil microbial communities and processes across biomes. Global Change Biology 18, 1918–1927.
Reich, P.B., Peterson, D.W., Wedin, D.A., Wrage, K., 2001. Fire and vegetation effects on productivity and nitrogen cycling across a forest–grassland continuum. Ecology 82, 1703–1719.
Rocca, J.D., Hall, E.K., Lennon, J.T., Evans, S.E., Waldrop, M.P., Cotner, J.B., Nemergut, D.R., Graham, E.B., Wallenstein, M.D., 2015. Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. The ISME Journal 9, 1693.
Salles, J.F., e Silva, M.C.P., Dini-Andreote, F., Dias, A.C., Guillaumaud, N., Poly, F., van Elsas, J.D., 2017. Successional patterns of key genes and processes involved in the microbial nitrogen cycle in a salt marsh chronosequence. Biogeochemistry 132, 185–201.
Shapleigh, J.P., 2013. Denitrifying Prokaryotes, the Prokaryotes. Springer, pp. 405–425.
Six, J., Frey, S., Thiet, R., Batten, K., 2006. Bacterial and fungal contributions to carbon sequestration in agroecosystems. Soil Science Society of America Journal 70, 555–569.
Smith, K., Dobbie, K., Ball, B., Bakken, L., Sitaula, B., Hansen, S., Brumme, R., Borken, W., Christensen, S., Priemé, A., 2000. Oxidation of atmospheric methane in Northern European soils, comparison with other ecosystems, and uncertainties in the global terrestrial sink. Global Change Biology 6, 791–803.
Song, W., Kim, M., Tripathi, B.M., Kim, H., Adams, J.M., 2016. Predictable communities of soil bacteria in relation to nutrient concentration and successional stage in a laboratory culture experiment. Environmental Microbiology 18 (6), 1740–1753.
Springael, D., Top, E.M., 2004. Horizontal gene transfer and microbial adaptation to xenobiotics: new types of mobile genetic elements and lessons from ecological studies. Trends in Microbiology 12, 53–58.
Sun, C.L., Relman, D.A., 2013. Microbiota's' little helpers': bacteriophages and antibiotic-associated responses in the gut microbiome. Genome Biology 14, 127.
Sun, S., Li, S., Avera, B.N., Strahm, B.D., Badgley, B.D., 2017. Soil bacterial and fungal communities show distinct recovery patterns during forest ecosystem restoration. Applied and Environmental Microbiology 83 (14), e00966-17 AEM. 00966-00917.
Top, E.M., Springael, D., 2003. The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds. Current Opinion in Biotechnology 14, 262–269.
Treseder, K.K., 2008. Nitrogen additions and microbial biomass: a meta-analysis of ecosystem studies. Ecology Letters 11, 1111–1120.
Wang, Z., Zhang, X.-X., Lu, X., Liu, B., Li, Y., Long, C., Li, A., 2014. Abundance and diversity of bacterial nitrifiers and denitrifiers and their functional genes in tannery wastewater treatment plants revealed by high-throughput sequencing. PLoS One 9, e113603.
Wei, W., Isobe, K., Nishizawa, T., Zhu, L., Shiratori, Y., Ohte, N., Koba, K., Otsuka, S., Senoo, K., 2015. Higher diversity and abundance of denitrifying microorganisms in environments than considered previously. The ISME Journal 9, 1954.
Weiss, S., Xu, Z.Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J.R., Vázquez-Baeza, Y., Birmingham, A., 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5, 27.
Yan, E.-R., Wang, X.-H., Guo, M., Zhong, Q., Zhou, W., Li, Y.-F., 2009. Temporal patterns of net soil N mineralization and nitrification through secondary succession in the subtropical forests of eastern China. Plant and Soil 320, 181–194.
Zak, D.R., Grigal, D.F., Gleeson, S., Tilman, D., 1990. Carbon and nitrogen cycling during old-field succession: constraints on plant and microbial biomass. Biogeochemistry 11, 111–129.
Zeng, J., Lou, K., Zhang, C.-J., Wang, J.-T., Hu, H.-W., Shen, J.-P., Zhang, L.-M., Han, L.-L., Zhang, T., Lin, Q., 2016. Primary succession of nitrogen cycling microbial communities along the deglaciated forelands of tianshan mountain, China. Frontiers in Microbiology 7.
Zerva, A., Mencuccini, M., 2005. Short-term effects of clearfelling on soil CO 2, CH 4, and N 2 O fluxes in a Sitka spruce plantation. Soil Biology and Biochemistry 37, 2025–2036.
Zhou, J., Deng, Y., Zhang, P., Xue, K., Liang, Y., Van Nostrand, J.D., Yang, Y., He, Z., Wu, L., Stahl, D.A., Hazen, T.C., Tiedje, J.M., Arkin, A.P., 2014. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. Proceedings of the National Academy of Sciences 111, E836–E845.
Zumsteg, A., Luster, J., Göransson, H., Smittenberg, R., Brunner, I., Bernasconi, S., Zeyer, J., Frey, B., 2012. Bacterial, archaeal and fungal succession in the forefield of a receding glacier. Microbial Ecology 63, 552–564.