Prokaryotic community shifts during soil formation on sands in the tundra zone

Alena Zhelezova1,*, Timofey Chernov1, Azida Tkhakakhova1, Natalya Xenofontova1,2, Mikhail Semenov1, Olga Kutovaya1

1 V.V. Dokuchaev Soil Science Institute, Department of Soil Biology and Biochemistry, Moscow, Russia,
2 Lomonosov Moscow State University, Department of Soil Science, Moscow, Russia

* zhelezova.ad@gmail.com

Abstract

A chronosequence approach, i.e., a comparison of spatially distinct plots with different stages of succession, is commonly used for studying microbial community dynamics during paedogenesis. The successional traits of prokaryotic communities following sand fixation processes have previously been characterized for arid and semi-arid regions, but they have not been considered for the tundra zone, where the environmental conditions are unfavourable for the establishment of complicated biocoenoses. In this research, we characterized the prokaryotic diversity and abundance of microbial genes found in a typical tundra and wooded tundra along a gradient of increasing vegetation—unfixed aeolian sand, semi-fixed surfaces with mosses and lichens, and mature soil under fully developed plant cover. Microbial communities from typical tundra and wooded tundra plots at three stages of sand fixation were compared using quantitative polymerase chain reaction (qPCR) and high-throughput sequencing of 16S rRNA gene libraries. The abundances of ribosomal genes increased gradually in both chronosequences, and a similar trend was observed for the functional genes related to the nitrogen cycle (nifH, bacterial amoA, nirK and nirS). The relative abundance of Planctomycetes increased, while those of Thaumarchaeota, Cyanobacteria and Chloroflexi decreased from unfixed sands to mature soils. According to β-diversity analysis, prokaryotic communities of unfixed sands were more heterogeneous compared to those of mature soils. Despite the differences in the plant cover of the two mature soils, the structural compositions of the prokaryotic communities were shaped in the same way. Thus, sand fixation in the tundra zone increases archaeal, bacterial and fungal abundances, shifts and unifies prokaryotic communities structure.

Introduction

For the investigation of microbial succession during soil-forming processes, a chronosequence approach, i.e., a comparison of spatially distinct plots of different ages, is commonly used. Currently, chronosequences of soil formation can be observed in areas with variable climatic conditions and on omnigenous parent material, such as glacial retreats [1–4], sand dunes [5–7],
volcanic rocks [8], or anthropogenic landscapes [9]. Changes in microbial community structure during the process of sand dune fixation have mostly been studied for arid and semi-arid regions [6,10,11] and for coastal environments [5]. Currently, successional traits during sand fixation in the cold climate of the tundra have received increasing attention. Biogeoecoenoses of Subarctic region play an important role in regulating the global carbon balance, but they are considered to be susceptible to consequences of climate change (e.g. an increase of mean annual temperatures), and have vulnerable vegetation cover [12]. The environmental conditions for soil formation in the tundra zone are specific, when the sandy substrates are depleted of nutrients, and the average temperature is unfavourable for the development of highly productive plant community.

While the succession of plant communities is relatively well studied, information on the prokaryotic community assemblage during soil formation is still lacking [10,13]. It is known that the first organisms to colonize parent rock are phototrophs, diazotrophs, chemolithotrophs and heterotrophs, whose taxonomic composition depends on the substrate properties [4,8,14]. Several bacterial phyla have been suggested to be associated with the initial stages of soil formation, mainly Bacteroidetes [3] and Cyanobacteria [15,16]. The prokaryotic community acts as the primary producer of organic matter and modifies the parent material for further colonization by plants. Available nitrogen is a limiting factor of plant growth, especially on lean substrates in cold environments [17–19]. Some prokaryotes (e.g. from phyla Cyanobacteria, Proteobacteria, Firmicutes, Actinobacteria) are able to perform nitrogen fixation, which leads to the accumulation of available nitrogen during inhabitation of barren substrates, such as rocks and sands [4]. Both archaeal and bacterial ammonia oxidizers produce nitrate (NO$_3^-$), which appears to be a crucial form of nitrogen for plants in the tundra zone [20]. Denitrification is a multi-step process of full or partial NO$_3^-$ reduction, which may lead to nitrogen losses through N$_2$ and N$_2$O emission [21,22]. Additionally, the presence of vegetation shapes prokaryotic community structure during soil formation [23]. In comparison to bacteria, fungi are less adapted to life on barren substrates and depend strongly on plants during the early stages of colonization [24].

The diversity of soil microorganisms changes during the process of soil formation; however, there is no distinct and universal pattern of prokaryotic diversity shifts with the successional stage of paedogenesis [2,3,23,25]. Previous studies have shown that at the earliest stages of soil formation after the retreat of glaciers (0–100 years), the bacterial diversity was relatively low, whereas it increased with the age of soil [2] or was the highest in middle-aged soils [3]. In contrast, in a longer timescale of ecosystem development (60–120 000 years), the diversity of the soil prokaryotic community decreased with the site age [25]. The patterns of prokaryotic diversity change among chronosequences of soil formation have been mostly studied for glacier retreats but not for soils formed on aeolian sand dunes.

The aim of this research was to reveal the traits of microbial community succession during sand fixation in the tundra zone. Two chronosequences of soil formation on aeolian sands with similar initial stages and different mature vegetation (typical tundra and wooded tundra) were compared. Taxonomic composition and diversity of the prokaryotic community, the abundances of bacterial, archaenal, and fungal ribosomal genes and functional genes related to the N cycle were estimated for three stages of sand fixation (unfixed sand—semi-fixed surface—mature soil). We hypothesized that 1) the ribosomal and functional genes abundances increases along the chronosequences, 2) $\alpha$-diversity of prokaryotic communities increases gradually with soil formation and plant colonization on sands, and 3) unfixed sands harbour similar prokaryotic community structures, while the communities in mature soils under the two vegetation types vary from each other.
Materials and methods

Sampling site description

Sand fixation chronosequences at two sites on the shores of the Pechora River (Northwestern Russia, Nenetsia region) were studied. This region is located in the southern tundra zone with a humid subarctic climate and an average annual temperature of -3.6 °C. The mean annual precipitation is 445 mm. For both sites, sampling was performed in August 2015 on three types of surfaces: 1 – unfixed aeolian sand, 2 – semi-fixed surface with mosses and lichens, and 3 – mature soil under developed plant cover (Fig 1). The two sites differed in the plant cover that developed on mature soil—typical tundra vegetation with subshrubs (Site I) and wooded tundra with rare trees and subshrubs (Site II).

The first site (Site I) was located on a flat sand hill, probably a moraine formation, with a height of approximately 30 m (67˚58′34.3″N, 52˚55′19.9″E, near Nelmin Nos). The areas of unfixed surface (sand with gravel and rare moss and grass shoots) were found on the hilltop. Presumably, the substrate on that surface was unfixed due to wind and snow erosion. The semi-fixed surface was covered by scanty vegetation: the cushion-like moss Racomitrium canescens, the lichen Stereocaulon paschale, rare subshrubs and other lichens. Vegetation on the mature soil was typical for the tundra zone: lichens (mostly Cladonia arbuscula and Flavocetraria nivalis), subshrubs (g. Empetrum, g. Arctostaphylos, g. Ledum), and f. Gramineae. The soil was classified as Arenosol in the WRB classification [26].

The second site (Site II) was located in a deflation basin with unfixed aeolian sand, which formed small dunes and was gradually covered by vegetation (67˚36′23.2″N, 53˚08′12.2″E, near Naryan-Mar). The unfixed surface was an aeolian sand without gravel. The semi-fixed surface was partly covered by shoots of moss (g. Polytrichum) and lichens (mostly Stereocaulon paschale). Vegetation on the mature soil consisted of various lichens, subshrubs (g. Empetrum, g. Arctostaphylos, Vaccinium vitis-idaea), grasses (Festuca rubra) and small trees (g. Juniperus, g. Betula). The soil was classified as Arenosol in the WRB classification [26] or Psammoozem on buried podzol.

For every surface type on each site, five samples were taken from depths of 1–5 cm that lacked plants, mosses and lichens. Sampling plots of different types were located on a transect...
with 3–5 m between each plot. For molecular analyses, samples were stored at -70 °C. The total organic carbon (TOC) and total nitrogen (TN) contents were estimated for the average sample from each plot using a Vario MACRO Cube CN-analyser (Elementar Analysensysteme GmbH, Germany).

**DNA extraction**

Total DNA was extracted from 0.5 g of frozen samples using the FastDNA SPIN kit for Soil (MP Biomedicals, USA) as recommended by the manufacturer. The homogenization step was performed with a Precellys 24 homogenizer (Bertin Technologies, France), program 5 (30 sec, 6500 rev. / min). DNA quality was estimated by electrophoresis in agarose gels (1% w/v in TAE) with further visual DNA detection using the Gel Doc XR+ System (Bio-Rad Laboratories, USA). DNA quantity was estimated by Qubit 3 Fluorometer (Thermo Fisher Scientific, USA) using Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, USA).

**Quantitative PCR analysis**

qPCR assays were used for quantitative estimation of ribosomal and N-cycle genes. To estimate the functional potential of microbial communities for N-fixation, ammonia-oxidation and denitrification, the abundance of genes encoding key enzymes of these processes (\textit{nifH}, \textit{amoA}, \textit{nirK} and \textit{nirS}, respectively) was measured [18]. 16S ribosomal genes of Bacteria and Archaea, the ITS region of Fungi and functional genes \textit{nifH}, bacterial \textit{amoA}, \textit{nirK} and \textit{nirS} were quantified using primer sets described in Table 1. All reactions were performed in a C1000 Thermal Cycler with the CFX96 Real-Time System (Bio-Rad Laboratories, USA). The qPCR mix contained 10 μl of 2X concentrated master mix for qPCR (SYBR Green Supermix (Bio-Rad Laboratories, USA) for the ITS region of Fungi, BioMaster HS-qPCR SYBR Blue (Biolabmix, Russia) for the other genes), 0.5–0.8 μM of each primer, and 1 μl of extracted soil DNA template in a total volume of 20 μl. Quantification of the initial gene copy abundance was performed in CFX Manager. PCR conditions for ribosomal genes were 3 min at 95 °C, followed by 49 cycles of 95 °C for 10 sec, 50 °C for 10 sec, and 72 °C for 20 sec. PCR conditions for N-cycle genes were 3 min at 95 °C, followed by 40 cycles of 95 °C for 20 sec, 54 °C for 20 sec, and 72 °C for 20 sec. To ensure qPCR specificity, melting curve analysis was performed (from 65 °C to 95 °C with an increment of 0.5 °C). Triplicate standard curves ranged from 10^3

| Target or process | Target gene | Primer name | Primer sequence (F, R) | Standard source | Reference |
|-------------------|-------------|-------------|------------------------|-----------------|-----------|
| **Total Bacteria** | 16S rRNA    | Eub338      | ACTCCTACGGGAGGCAGCAG    | ATTACCGCGGCTGCTGG | Esherichia coli [27] |
|                   |             | Eub518      |                        |                  |           |
| **Total Archaea** | 16S rRNA    | 915f        | AGAA TTGGC GGGGG AGCAC | GCCAT GCACC WCCTC T | strain FG-07 Halobacterium salinarum [28] |
|                   |             | 1059r       |                        |                  |           |
| **Total Fungi**   | ITS region  | ITS1f       | TCC GTA GGT GAA CCT GCG G GCG TGC GTT CTT CAT CG | Saccharomyces cerevisae Meyen 1B-D1606 [27] |
|                   |             | 5.8s        |                        |                  |           |
| **N-fixation**    | \textit{nifH} | PolF        | TGC GAY CCS AAR GCB GAC TC ATS GCC ATC ATY TCR CCG GA | Sinorhizobium meliloti [29] |
|                   |             | PolR        |                        |                  |           |
| **nitrification** | Bacterial \textit{amoA} | amoA-1F   | GGGGTTCCTACTGGTGGTGTTGTTG | CCCCCTGGAAGAGGCTCTTTC | Standard was generated by PCR amplification of \textit{amoA} genes from extracted DNA from mature soil of Site 1 [30] |
|                   |             | amoA-2R    |                        |                  |           |
| **denitrification** | \textit{nirK} | nirK876    | ATY GGC GGV CAY GGC GA GGC TGC ATC AGR TTR TGG TT | Sinorhizobium meliloti [30] |
|                   |             | nirK1040   |                        |                  |           |
|                   | \textit{nirS} | cd3af      | GTSACC GSAAGGARACS GG  | GASTTGCGRTRGGTYCTTGA | Pseudomonas sp [31,32] |
|                   |             | R3cd       |                        |                  |           |

https://doi.org/10.1371/journal.pone.0206777.t001
to $10^8$ gene copy number/μl. Standards were made by purifying PCR products and quantifying the concentration by Qubit fluorometer 2 (Thermo Fisher Scientific, USA). Reference organisms (except amoA gene) for the construction of standard curves for PCR products are described in the Table 1. Efficiencies of qPCR were 82–101% and coefficients of determination were $R^2 > 0.90$ for all standard curves.

**Sequencing of 16S rRNA gene libraries**

High-throughput sequencing of the 16S rRNA gene libraries was performed for 5 replicates of each studied sample. The purified DNA isolates were amplified with universal multiplex primers F515 (5′-GTGCCAGCMGCCGCGGTAA-3′) and R806 (5′-GACTACVSGGGGTATCTAAT-3′) [33] targeting variable regions V3–V4 of bacterial and archaeal 16S rRNA genes. PCR was carried out in a 15 μl reaction mixture containing 0.5–1 units of Phusion Hot Start II High-Fidelity polymerase and 1X Phusion buffer (Thermo Fisher Scientific, USA), 5 pM of forward and reverse primers, 10 ng of DNA matrix and 2 nM of each dNTP (Thermo Fisher Scientific, USA). The mixture was denatured at 94 °C for 1 min, followed by 35 cycles of 94 °C for 30 sec, 50 °C for 30 sec, and 72 °C for 30 sec. The final elongation was carried out at 72 °C for 3 min. PCR products were purified according to the recommended Illumina technique using AM Pure XP (Beckman Coulter, USA). Further preparation of the 16S rRNA gene libraries was carried out as described in the MiSeq Reagent Kit Preparation Guide (Illumina, USA). Sequencing of 16S rRNA gene amplicons was carried out on an Illumina MiSeq platform using MiSeq Reagent Kit v3 (600 cycles) with forward and reverse reading. The raw data is deposited in NCBI database (BioProject ID: PRJNA497067).

**Processing of 16S rRNA gene data**

Sequencing data were processed using QIIME v.1.9.1 [34] and Trimmomatic [35]. Sequence pairs with both forward and reverse reads of at least 180 nucleotides were merged using the fastq-join algorithm. Trimming, i.e. filtering of sequences according to the reading quality parameters was performed using the Trimmomatic program [35], so that the quality of 4 adjacent nucleotides was not lower than 16. Operational taxonomic units (OTU) picking based on 97% nucleotide similarity was performed in the QIIME environment. Chimeras were filtered using VSEARCH algorithm [36]. Reference sequences for OTU selection as well as taxonomic affiliation were obtained from SILVA database version 128, 2017 (https://www.arb-silva.de/download/archive/qiime). Singletons (OTUs containing only one sequence) and 16S rRNA sequences of chloroplasts and mitochondria were removed.

**Statistical and sequence analyses**

Statistical analysis of gene abundance data was performed in Microsoft Excel and STATISTICA 10.0. A multiple t-test was performed to test for significant ($p<0.05$) differences between gene abundances in three plot types of each chronosequence. Pearson correlation test was performed to check correlations between substrate chemical properties and gene abundances.

Several indices were used for the estimation of total diversity of the studied prokaryotic communities ($\alpha$-diversity). The Shannon index was calculated ($H = \sum pi \ln pi$, where $pi$ is the relative abundance of species $i$ in the community). The Chao1 index and the phylogenetic diversity whole tree metric were calculated for characterization of the real number of OTUs in the prokaryotic community [37,38]; they were compared with the total number of observed OTUs. Data were normalized to 4090 sequences per sample.

The analysis of structural differences between prokaryotic communities ($\beta$-diversity) was performed using binary metrics of similarity—weighted UniFrac, unweighted UniFrac and
Bray-Curtis metrics [39]. Based on weighted UniFrac distances, non-metric multidimensional scaling (NDMS) was carried out to construct diagrams of similarity in prokaryotic community structures. The significance of the differences between structures of prokaryotic communities on three stages of soil formation was estimated using ANOSIM (Analysis of similarities) script in QIIME (999 permutations).

**Results**

**Chemical properties of substrates on the studied plots**

Both TOC and TN contents increased in correspondence with the stage of soil formation. For both sampling sites, unfixed sands and semi-fixed surfaces had extremely low organic carbon (0–0.18%) and nitrogen (0.02–0.04%) contents, while mature soils had much higher amounts of C and N (Table 2). The difference in percentage of TOC between the studied plots was higher than the difference in N content. The amount of DNA recovered from samples increased along two chronosequences.

**Ribosomal and N-cycle gene abundances in the two chronosequences**

The gradual increase in ribosomal gene copy numbers was revealed in both chronosequences from the unfixed sand to the mature soil. At Site I, there was a statistically significant increase (p<0.001) of one order of magnitude in bacterial and fungal gene copy numbers per gram of substrate, while archaeal gene copy number increased by two orders of magnitude (Fig 2). The gene abundances in the samples from the semi-fixed surfaces were intermediate between those of the unfixed sand and the mature soil. An increase of two orders of magnitude in the number of all ribosomal genes was also observed for Site II; however, mean values of fungal gene copy numbers in the semi-fixed surface and in the mature soil did not significantly differ from each other (Fig 2, S1 Table).

Similar trends were observed for the distribution of functional genes along the chronosequences (Fig 3). For both sites, there was a two-order increase in the amount of all N-cycle genes from the unfixed sand to the mature soil. At Site I, the semi-fixed surface was more similar to the unfixed surface; at Site II, conversely, there was stronger similarity between the semi-fixed surface and the mature soil. The nirK gene, which is associated with denitrification, was the most abundant among the investigated functional genes in all samples, while amoA genes (associated with nitrification) were least abundant.

Bacterial gene abundance in the substrate correlated with TOC percentage (p<0.05), while archaeal gene abundance correlated with TN (p<0.05), and fungal gene abundance correlated with both TOC and TN (Table 3). All functional genes related to the nitrogen cycle were significantly correlated with TN (p<0.01 for nirK and nirS, p<0.005 for nifH and amoA).

**Prokaryotic community structure among the chronosequences**

In total, 261 161 sequences of the 16S rRNA gene were obtained (from 2458 to 21 196 sequences per sample) with a mean length of 292 bp. One sample was excluded from the analysis of α-diversity due to a low number of sequences.

Phyla Proteobacteria and Acidobacteria were predominant in all samples (up to 35% of relative abundance) (Fig 4). The taxonomic structure on the phylum level was similar for the prokaryotic communities in the two mature soils under different vegetation types. The comparison of abundances of different phyla showed that the relatively high abundances of Thaumarchaeota (up to 7% on Site I), Chloroflexi (up to 12%) and Cyanobacteria (up to 14% on Site II) were associated with unfixed sands, while Planctomycetes was more abundant in the
The prokaryotic community structure of semi-fixed surfaces was intermediate between those of unfixed sands and mature soils. *Chloroflexi* was relatively more abundant in all samples of unfixed sands and semi-fixed surfaces. Genera *Chamaesiphon* (up to 9% relative abundance), *Crinalium*, *Leptolyngbya* and *Stigonema* belonging to phylum *Cyanobacteria* were most abundant in the unfixed sand from Site II. There was no significant correlations found between relative abundance of phyla and TOC and TN amounts, except *Firmicutes* (S2 Table).

### The diversity of prokaryotic communities among the chronosequences

The highest prokaryotic α-diversity was found in mature soil from Site I (Table 4). Prokaryotic diversity indices for unfixed sand and the semi-fixed surface at Site I did not differ significantly.

![Ribosomal gene copy number in the plots of sites I and II. US indicates unfixed sand, SF—semi-fixed surface, MS—mature soil. The data are shown as means (n = 5). Error bars represent standard deviations.](https://doi.org/10.1371/journal.pone.0206777.g002)
Increased α-diversity was observed among the chronosequence at Site I, while no significant difference was revealed for prokaryotic communities at Site II due to high variation between indices for samples from each plot.

The shift in prokaryotic community composition during the process of sand fixation was observed in both chronosequences (Fig 5). Both unweighted and weighted UniFrac analyses showed similar community compositions of the two mature soils and semi-fixed surfaces, while the sand samples formed separate clusters and differed from each other. In weighted UniFrac metrics, prokaryotic communities of unfixed sands were separated from those of semi-fixed surfaces and mature soils. Bray-Curtis metrics also showed that prokaryotic communities of the two sands were significantly different from the other samples. For all metrics (Bray Curtis, Weighted UniFrac, Unweighted UniFrac) the differences between prokaryotic communities on

![Functional gene copy number in the plots of sites I and II.](https://doi.org/10.1371/journal.pone.0206777.g003)

Table 3. Significant correlations (p<0.01, *—p<0.005) between total organic carbon (TOC), nitrogen (TON) percentage and gene abundances.

| Targeted gene amplicon | TOC, % | TN, % |
|------------------------|--------|-------|
| 16S rRNA (bacterial)   | n.s.   | n.s.  |
| 16S rRNA (archaeal)    | n.s.   | n.s.  |
| ITS region (fungal)    | 0.94   | 0.94  |
| nifH                   | 0.96*  | 0.99* |
| amoA (bacterial)       | 0.96*  | 0.99* |
| nirK                   | n.s.   | 0.94  |
| nirS                   | n.s.   | 0.92  |

![Microbiome shifts during soil formation in tundra](https://doi.org/10.1371/journal.pone.0206777.t003)
three types of plots were significant (p < 0.01) in both chronosequences. According to R test, the difference between soil formation stages is more pronounced on the Site II (with wooded tundra vegetation on the mature soil plot) comparing to the Site I (S3 Table).

### Discussion

**Quantitative analysis of ribosomal and N-cycle genes in the two chronosequences**

The observed ribosomal gene copy numbers both in semi-fixed surfaces and in mature soils correspond with the previously obtained data on the microbial population abundance for soils.
and soil-like substrates in northern latitudes [40,41]. The higher abundance of bacterial ribosomal genes in comparison to the abundance of fungal genes in all samples can be explained by low diversity of plant communities at the studied plots; as previously shown in other studies, Fungi are more dependent on the plant communities of barren substrates than Bacteria [24]. However, the abundance of genes increased from unfixed sand to mature soil in both chronosequences, which can be an effect of plant community development and consequently higher available organic carbon and nitrogen contents. Available organic matter in soil is known to be a limiting factor of microbial community development [42], and a correlation between the quantity of ribosomal genes and soil organic carbon content was previously observed for other soils [43].

The abundances of all functional genes associated with transformation of nitrogen also increased in both chronosequences from unfixed sand to mature soil and correlated with total nitrogen content. In other studies, the abundance of nifH genes related to N\(_2\) fixation was found to be correlated with total nitrogen content, nitrate and ammonium concentrations [44,45]. This finding is consistent with the study of metagenomes of Arctic tundra soil, where N-assimilation genes were present in all bacterial genomes in microbiomes of different types of polygonal landscapes [46]. Another study showed that in soils formed on glacial retreats, the nitrogen fixation rates significantly increased during the first 4–5 years of succession [15]. Thus, nitrogen-fixing bacteria are important for soil microbial community assemblage and functioning, especially in the tundra zone characterized by scarce vegetation and low nitrogen content. The lowest abundance of the bacterial amoA gene among all functional genes studied can be explained by the low amount of organic nitrogen in all samples. Bacterial amoA gene abundance in soil is known to be related to the available ammonia concentration [47]. In all samples, nirK gene abundance was the highest in comparison to other N-cycle genes. Genes

---

**Fig 5. Beta-diversity indices of microbial communities in samples.** I, II are the site numbers, US indicates unfixed sand samples, SF – semi-fixed surfaces, MS – mature soil samples.

https://doi.org/10.1371/journal.pone.0206777.g005
associated with denitrification (nirK and nirS) were previously found to be more abundant than nifH and amoA in other soils [44]. The abundances of the two nitrite reductase-encoding genes (nirK and nirS) observed in this study were disproportionate. Similar trend with higher abundance of nirK genes comparing to nirS was previously observed in Arctic soils [12].

According to previous studies, nirK (copper nitrite reductase) is more widespread in terrestrial ecosystems, while nirS (cytochrome cd1 nitrite reductase) is more abundant in marine environments [48–50].

Thus, the obtained data for both Sites supported the hypothesis of increased gene abundance among the chronosequences. The similar dynamics of ribosomal and functional gene abundance in the two chronosequences can be explained by the congruent patterns of increased total organic carbon and total nitrogen from unfixed sands to mature soils on the two sites. Gene copy number indirectly indicates the biomass of different functional and taxonomic groups in soil microbial communities, which increases during primary succession on different types of barren substrates [7].

Changes of prokaryotic community structure during soil formation

The prokaryotic communities of all samples were dominated by phyla Acidobacteria and Proteobacteria, which was previously observed for other soils of the tundra zone [46,51].

Phylum Thaumarchaeota was found to be relatively more abundant in unfixed sands of both sites. All OTUs belonging to Thaumarchaeota were uncultivable Archaea and were previously observed in other terrestrial environments. Among all Archaeal phyla, Thaumarchaeota are known to be predominate in Arctic and Antarctic soils [51]. We suggest that the relative abundance of Thaumarchaeota, but not their absolute number, decreased from unfixed sands to mature soils because archaeal gene abundances in the unfixed sands were a hundred-fold lower than in the mature soils. It is also possible that archaeal OTUs from mature soils were not presented in the database and were ranked as unclassified.

Family Ktedonobacteraceae belonging to phylum Chloroflexi that were predominant in unfixed sands are known to be negatively correlated with organic matter content in deforested soil [52]. This family is mostly represented by uncultivable genera; its cultivable representatives are filamentous, aerobic and mesophilic [53]. Ktedonobacteraceae was previously found in young soils [23,54,55]. In this study, the relative abundance of Chloroflexi decreased with succession development, what is in accordance with previous findings [23].

Representatives of phylum Cyanobacteria were more abundant in the samples of unfixed sand from Site II, which can be explained by low levels of insolation of sand on Site I due to gravel cover. Cyanobacteria are known as free-living phototrophs capable of nitrogen fixation, especially in extreme environments [16,51,56]. Representatives of this phylum could be the primary producers of organic matter in unfixed sands due to the lack of organic carbon and nitrogen. A decrease in Cyanobacteria abundance and number of observed OTUs belonging to this phylum with soil age was previously observed for soils formed by glacial isostatic adjustment in Fennoscandia [57]. Genus Leptolyngbya was found on plots of unfixed sand on Site II. It is known as a producer of adhesive extra-cellular polysaccharides and organic acids that can degrade rock [58]. Thus, all these taxa inhabit barren substrates, and their active presence in the community can be considered an indicator of the primary stage of development of microbial succession.

Diversity changes among the chronosequences

The gradual increase of prokaryotic α-diversity from initial stages of sand fixation to mature soils was expected for both chronosequences. However, prokaryotic α-diversity increased
from the unfixed sands to the mature soil on Site I, while prokaryotic communities of all samples on Site II did not follow the same pattern, and their α-diversities did not change with the successional stage. The obtained results partly contradict previously discovered trends of incremental growth of prokaryotic α-diversity during revegetation on moving dunes [10]. However, some studies reported the highest prokaryotic diversity at the early stages of soil formation [59,60]. Although the microbial biomass on barren substrates was relatively low, the high diversity of the unfixed sand prokaryotic community can be explained by the variety of necessary adaptations to harsh environmental conditions.

The prokaryotic community structures in samples of the two mature soils appeared to be very similar comparing to those in the unfixed sand samples. The estimation of β-diversity showed the same pattern of unevenness of prokaryotic communities in the unfixed sand samples. This observed dissimilarity could be a consequence of random propagule input in unfixed sands, while the developed vegetation on both mature soils allowed the formation of more stable prokaryotic communities. Another possible explanation of prokaryotic community differences in unfixed sands could be that the diversity depends on both richness and evenness [61–63], and microbiomes in unfixed sands were less even, than in mature soils.

Conclusions
Using a chronosequence approach, we found the expected trends in microbial populations: increased microbial community abundance and change of prokaryotic community structure from unfixed sands to mature soils. The highest prokaryotic diversity and abundance, as well as the amount of microorganisms involved in the nitrogen cycle, were revealed in mature soil under developed plant cover. However, the prokaryotic diversity during soil formation increased slightly, with the minimum values found in sand under pioneer vegetation (intermediate stages of succession). In contrast with our predictions, the analysis of β-diversity shown that prokaryotic communities under the unfixed sands were more dissimilar, than under vegetation. Therefore, plant colonization of aeolian sands in tundra multiplies and unifies the prokaryotic communities.

Supporting information
S1 Table. Results of t-test for independent samples (pairwise comparisons) of gene abundances. Significant difference (p<0.05) is marked red.
(XLSX)

S2 Table. Correlations between total organic carbon (TOC), nitrogen (TON) percentage and phyla abundances. Significant correlations (p<0.05) are marked red.
(XLSX)

S3 Table. Results of ANOSIM. R values and p values calculated for Weighted UniFrac, Unweighted UniFrac and Bray-Curtis metrics.
(XLSX)

S1 Dataset. Minimal data set Figs 2 and 3.
(XLSX)

S2 Dataset. Minimal data set Fig 5.
(XLSX)
Acknowledgments

We would like to thank Sergei Khokhlov and Dmitry Konuyshkov for their support. Experiments were performed using the equipment of the Collective Use Center “Functions and Properties of Soils and Soil Cover” of the Dokuchaev Institute. High-throughput sequencing was performed using the Illumina MiSeq platform of the Common Use Center of Saint Petersburg State University.

Author Contributions

Conceptualization: Timofey Chernov, Mikhail Semenov.

Data curation: Alena Zhelezova, Timofey Chernov, Natalya Xenofontova.

Formal analysis: Alena Zhelezova.

Funding acquisition: Timofey Chernov.

Investigation: Alena Zhelezova, Azida Tkhakakhova, Natalya Xenofontova.

Methodology: Alena Zhelezova, Timofey Chernov.

Project administration: Azida Tkhakakhova, Olga Kutovaya.

Resources: Olga Kutovaya.

Supervision: Olga Kutovaya.

Validation: Alena Zhelezova, Timofey Chernov, Mikhail Semenov.

Visualization: Timofey Chernov.

Writing – original draft: Alena Zhelezova.

Writing – review & editing: Alena Zhelezova, Timofey Chernov, Mikhail Semenov.

References

1. Esposito A, Ciccazzo S, Borruso L, Zerbe S, Daffonchio D, Brusetti L. A three-scale analysis of bacterial communities involved in rocks colonization and soil formation in high mountain environments. Curr Microbiol. 2013; 67: 472–479. https://doi.org/10.1007/s00284-013-0391-9 PMID: 23712376

2. Wu X, Zhang W, Liu G, Yang X, Hu P, Chen T, et al. Bacterial diversity in the foreland of the Tianshan No. 1 glacier, China. Environ Res Lett. 2012; 7: 014038. https://doi.org/10.1088/1748-9326/7/1/014038

3. Sun H, Wu Y, Zhou J, Bing H. Variations of bacterial and fungal communities along a primary successional chronosequence in the Hailuogou glacier retreat area (Gongga Mountain, SW China). J Mt Sci. 2016; 13: 1621–1631. https://doi.org/10.1007/s11629-015-3570-2

4. Ciccazzo S, Esposito A, Borruso L, Brusetti L. Microbial communities and primary succession in high altitude mountain environments. Ann Microbiol. 2016; 66: 43–60. https://doi.org/10.1007/s13213-015-1130-1

5. Turner BL, Laliberté E. Soil Development and Nutrient Availability Along a 2 Million-Year Coastal Dune Chronosequence Under Species-Rich Mediterranean Shrubland in Southwestern Australia. Ecosystems. 2015; 18: 287–309. https://doi.org/10.1007/s10021-014-9830-0

6. Liu L, Liu Y, Hui R, Xie M. Recovery of microbial community structure of biological soil crusts in successional stages of Shapotou desert revegetation, northwest China. Soil Biol Biochem. Elsevier Ltd; 2017; 107: 125–128. https://doi.org/10.1016/j.soilbio.2016.12.030

7. Schaub I, Baum C, Schumann R, Karsten U. Effects of an Early Successional Biological Soil Crust from a Temperate Coastal Sand Dune (NE Germany) on Soil Elemental Stoichiometry and Phosphatase Activity. Microb Ecol. Microbial Ecology; 2018; 1–13. https://doi.org/10.1007/s00248-018-1220-2 PMID: 29926147

8. Kelly LC, Cockell CS, Thorsteinsson T, Marteinsson V, Stevenson J. Pioneer Microbial Communities of the Fimmvorthuals Lava Flow, Eyjafjallajokull, Iceland. Microb Ecol. 2014; 68: 504–518. https://doi.org/10.1007/s00248-014-0432-3
9. Zhan J, Sun Q. Diversity of free-living nitrogen-fixing microorganisms in the rhizosphere and non-rhizosphere of pioneer plants growing on wastelands of copper mine tailings. Microbiol Res. The Research Centre for Eco-Environmental Sciences, Chinese Academy of Sciences; 2012; 167: 157–165. https://doi.org/10.1016/j.mires.2011.05.006 PMID: 21665448

10. Cao C, Zhang Y, Cui Z, Feng S, Wang T, Ren Q. Soil bacterial community responses to revegetation of moving sand dune in semi-arid grassland. Appl Microbiol Biotechnol. Applied Microbiology and Biotechnology; 2017; 101: 6217–6228. https://doi.org/10.1007/s00253-017-8336-z PMID: 28567480

11. Yu J, Unc A, Zhang X, Steinberger Y. Responses of the soil microbial catabolic profile and diversity to vegetation rehabilitation in degraded semiarid grassland. Appl Soil Ecol. Elsevier B.V.; 2016; 101: 124–131. https://doi.org/10.1016/j.apsoil.2016.01.022

12. Blaud A, Lerch TZ, Phoenix GK, Osborn AM. Arctic soil microbial diversity in a changing world. Res Microbiol. Elsevier Masson SAS; 2015; 166: 796–813. https://doi.org/10.1016/j.resmic.2015.07.013 PMID: 26275598

13. Walker LR, Wardle DA, Bardgett RD, Clarkson BD. The use of chronosequences in studies of ecological succession and soil development. J Ecol. 2010; 98: 725–736. https://doi.org/10.1111/j.1365-2745.2010.01664.x

14. Nemergut DR, Anderson SP, Cleveland CC, Martin AP, Miller AE, Seimom A, et al. Microbial community succession in an unvegetated, recently deglaciated soil. Microb Ecol. 2007; 53: 110–122. https://doi.org/10.1007/s00248-006-9144-7 PMID: 17186150

15. Schmidt S., Reed SC, Nemergut DR, Stuart Grandy A, Cleveland CC, Weintraub MN, et al. The earliest stages of ecosystem succession in high-elevation (5000 metres above sea level), recently deglaciated soils. Proc R Soc B Biol Sci. 2008; 275: 2793–2802. https://doi.org/10.1098/rspb.2008.0808 PMID: 18756777

16. Chriasmas NAM, Anesio AM, Sánchez-Baracaldo P. The future of genomics in polar and alpine cyanobacteria. FEMS Microbiol Ecol. 2018; 1–10. https://doi.org/10.1093/femsec/fiy032 PMID: 29506259

17. Glanville HC, Hill PW, MacCarone LD, Golyshin PN, Murphy D V., Jones DL. Temperature and water controls on vegetation emergence, microbial dynamics, and soil carbon and nitrogen fluxes in a high Arctic tundra ecosystem. Funct Ecol. 2012; 26: 1366–1380. https://doi.org/10.1111/j.1365-2435.2012.02056.x

18. Chapin FS, Matson P a, Vitousek PM. Principles of Terrestrial Ecosystem Ecology [Internet]. Springer. New York, NY: Springer New York; 2011. https://doi.org/10.1007/978-1-4419-9504-9

19. Castle SC, Sullivan BW, Knelman J, Hood E, Nemergut DR, Schmidt SK, et al. Nutrient limitation of soil microbial activity during the earliest stages of ecosystem development. Oecologia. Springer Berlin Heidel.; 2017; https://doi.org/10.1007/s00442-017-3965-6 PMID: 28993721

20. Liu X-Y, Koba K, Koyama LA, Hobbie SE, Weiss MS, Inagaki Y, et al. Nitrate is an important nitrogen source for Arctic tundra plants. Proc Natl Acad Sci. 2018; 115: 3398–3403. https://doi.org/10.1073/pnas.1715382115 PMID: 29540568

21. Henry S, Baudoin E, López-Gutiérrez JC, Martin-Laurent F, Brauman A, Philippot L. Quantification of denitrifying bacteria in soils by nirK gene targeted real-time PCR. J Microbiol Methods. 2004; 59: 327–335. https://doi.org/10.1016/j.mimet.2004.07.002 PMID: 15488276

22. Madsen EL. Microorganisms and their roles in fundamental biogeochemical cycles. Curr Opin Biotechnol. Elsevier Ltd; 2011; 22: 456–464. https://doi.org/10.1016/j.copbio.2011.01.008 PMID: 21333523

23. Brown SP, Jumpponen A. Contrasting primary successional trajectories of fungi and bacteria in retreating glacier soils. Mol Ecol. 2014; 23: 481–497. https://doi.org/10.1111/mec.12487 PMID: 24112459

24. Schmidt SK, Nemergut DR, Darcy JL, Lynch R. Do bacterial and fungal communities assemble differently during primary succession? Mol Ecol. 2014; 23: 254–258. https://doi.org/10.1111/mec.12589 PMID: 26010467

25. Jangid K, Whitman WB, Condron LM, Turner BL, Williams MA. Soil bacterial community succession during long-term ecosystem development. Mol Ecol. 2013; 22: 3415–3424. https://doi.org/10.1111/mec.12325 PMID: 24624422

26. ISSS. World reference base for soil resources 2014 [Internet]. 2014. ISSN 0532-0488

27. Fierer N, Jackson JA, Vilgalys R, Jackson RB. Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. Appl Environ Microbiol. 2005; 71: 4117. https://doi.org/10.1128/AEM.71.7.4117-4120.2005 PMID: 16000830

28. Yu Y, Lee C, Hwang S. Analysis of community structures in anaerobic processes using a quantitative real-time PCR method. Water Sci Technol. 2005; 52: 85–91.

29. Poly F, Ranjard L, Nazaret S, Gourbière F, Jocteur Monrozier L. Comparison of nifH Gene Pools in Soils and Soil Microenvironments with Contrasting Properties. Appl Environ Microbiol. 2001; 67: 2255–2262. https://doi.org/10.1128/AEM.71.5.2255-2262.2001 PMID: 11319109
Microbiome shifts during soil formation in tundra

30. Hallin S, Jones CM, Schloter M, Philippot L. Relationship between N-cycling communities and ecosystem functioning in a 50-year-old fertilization experiment. Isme J. Nature Publishing Group; 2009; 3: 597–605. https://doi.org/10.1038/ismej.2008.128 PMID: 19148144

31. Michotey V, Méjean V, Bonin P. Comparison of methods for quantification of cytochrome cd1-denitrifying bacteria in environmental marine samples. Appl Environ Microbiol. 2000; 66: 1564–1571. https://doi.org/10.1128/AEM.66.4.1564-1571.2000 PMID: 10742243

32. Throbäck IN, Enwall K, Jarvis A, Hallin S. Reassessing PCR primers targeting nirS, nirK and nosZ genes for community surveys of denitrifying bacteria with DGGE. FEMS Microbiol Ecol. 2004; 49: 401–417. https://doi.org/10.1016/j.femsec.2004.04.011 PMID: 19712290

33. Bates ST, Berg-Lyons D, Caporaso JG, Walters WA, Knight R, Fierer N. Examining the global distribution of dominant archaeal populations in soil. ISME J. Nature Publishing Group; 2011; 5: 908–917. https://doi.org/10.1038/ismej.2010.171 PMID: 21085198

34. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010; 7: 335–336. https://doi.org/10.1038/nmeth.f.303 PMID: 20383131

35. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

36. Quince C, Nichols B, Rognes T, Flouri T, Mahe F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016; 4: e2584. https://doi.org/10.7717/peerj.2584 PMID: 27781170

37. Chao A. Nonparametric Estimation of the Number of Classes in a Population Author. Scandinavian J Stat. 1984; 11: 265–270.

38. Colwell RK, Coddington JA. Estimating Terrestrial Biodiversity through Extrapolation. Philos Trans R Soc B Biol Sci. 1994; 345: 101–118. https://doi.org/10.1098/rstb.1994.0091 PMID: 7972351

39. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: An effective distance metric for microbial community comparison. ISME J. Nature Publishing Group; 2011; 5: 169–172. https://doi.org/10.1038/ismej.2010.133 PMID: 20827291

40. Buckeridge KM, Banerjee S, Siciliano SD, Grogan P. The seasonal pattern of soil microbial community structure in mesic low arctic tundra. Soil Biol Biochem. Elsevier Ltd; 2013; 65: 338–347. https://doi.org/10.1016/j.soilbio.2013.06.012

41. Ren J, Song C, Hou A, Song Y, Zhu X, Cagle GA. Shifts in soil bacterial and archaeal communities during freeze-thaw cycles in a seasonal frozen marsh, Northeast China. Sci Total Environ. Elsevier B.V.; 2018; 625: 782–791. https://doi.org/10.1016/j.scitotenv.2017.12.309 PMID: 29306166

42. Kuzyakov Y, Blagodatskaya E. Microbial hotspots and hot moments in soil: Concept & review. Soil Biol Biochem. Elsevier Ltd; 2015; 83: 184–199. https://doi.org/10.1016/j.soilbio.2015.01.025

43. Semenov M V., Chernov TI, Tkhakakhova AK, Zhelezova AD, Ivanova EA, Kolganova T V., et al. Distribution of prokaryotic communities throughout the Chernozem profiles under different land uses for over a century. Appl Soil Ecol. Elsevier; 2018; 127: 8–18. https://doi.org/10.1016/j.apsoil.2018.03.002

44. Regan K, Stempfhuber B, Schloter M, Rasche F, Prati D, Philippot L, et al. Spatial and temporal dynamics of nitrogen fixing, nitrifying and denitrifying microbes in an unfertilized grassland soil. Soil Biol Biochem. Elsevier Ltd; 2017; 109: 214–226. https://doi.org/10.1016/j.soilbio.2016.11.011

45. Ma W, Jiang S, Assemien F, Qin M, Ma B, Xie Z, et al. Response of microbial functional groups involved in soil N cycle to N, P and NP fertilization in Tibetan alpine meadows. Soil Biol Biochem. 2016; 101: 195–206. https://doi.org/10.1016/j.soilbio.2016.07.023

46. Taş N, Prestat E, Wang S, Wu Y, Ulrich C, Knaefsey T, et al. Landscape topography structures the soil microbiome in arctic polygonal tundra. Nat Commun. 2018; 9: 777. https://doi.org/10.1038/s41467-018-03089-z PMID: 29472560

47. Prosser JL, Nicol GW. Archaeal and bacterial ammonia-oxidisers in soil: the quest for niche specialisation and differentiation. Trends Microbiol. Elsevier Current Trends; 2012; 20: 523–531. https://doi.org/10.1016/j.tim.2012.08.001 PMID: 22959489

48. Jones CM, Hallin S. Ecological and evolutionary factors underlying global and local assembly of denitrifier communities. ISME J. Nature Publishing Group; 2010; 4: 633–641. https://doi.org/10.1038/ismej.2009.152 PMID: 20090785

49. Braker G, Fessefeldt A, Witzel KP. Development of PCR primer systems for amplification of nitrite reductase genes (nirK and nirS) to detect denitrifying bacteria in environmental samples. Appl Environ Microbiol. 1998; 64: 3769–75. Available: http://aem.asm.org/cgi/content/abstract/64/10/3769 PMID: 9758798

50. Yang Y, Zhao J, Jiang Y, Hu Y, Zhang M, Zeng Z. Response of bacteria harboring nirS and nirK genes to different N fertilization rates in an alkaline northern Chinese soil. Eur J Soil Biol. Elsevier Masson SAS; 2017; 82: 1–9. https://doi.org/10.1016/ejsoilb.2017.05.006
51. Makhalanyane TP, Van Goethem MW, Cowan DA. Microbial diversity and functional capacity in polar soils. Curr Opin Biotechnol. Elsevier Ltd; 2016; 38: 159–166. https://doi.org/10.1016/j.copbio.2016.01.011 PMID: 26921734

52. Navarrete AA, Tsai SM, Mendes LW, Faust K, De Hollander M, Cassman NA, et al. Soil microbiome responses to the short-term effects of Amazonian deforestation. Mol Ecol. 2015; 24: 2433–2448. https://doi.org/10.1111/mec.13172 PMID: 25809788

53. Yabe S, Sakai Y, Abe K, Yokota A, Také A, Matsumoto A, et al. Dictyobacter aurantiacus gen. nov., sp. nov., a member of the family Ktedonobacteraceae, isolated from soil, and emended description of the genus thermosporothrix. Int J Syst Evol Microbiol. 2017; 67: 2615–2621. https://doi.org/10.1099/ijsem.0.001985 PMID: 28758628

54. Rughoft S, Herrmann M, Lazar CS, Cesarz S, Levick SR, Trumbore SE, et al. Community composition and abundance of bacterial, archaeal and nitrifying populations in savanna soils on contrasting bedrock material in Kruger National Park, South Africa. Front Microbiol. 2016; 7. https://doi.org/10.3389/fmicb.2016.01638 PMID: 27807431

55. Li J, Wang JT, Hu HW, Ma YB, Zhang LM, He JZ. Copper pollution decreases the resistance of soil microbial community to subsequent dry-rewetting disturbance. J Environ Sci (China). Elsevier B.V.; 2016; 39: 155–164. https://doi.org/10.1016/j.jes.2015.10.009 PMID: 26899654

56. Levy-Booth DJ, Prescott CE, Grayston SJ. Microbial functional genes involved in nitrogen fixation, nitrification and denitrification in forest ecosystems. Soil Biol Biochem. 2014; 75: 11–25. https://doi.org/10.1016/j.soilbio.2014.03.021

57. Yarwood SA, Högberg MN. Soil bacteria and archaea change rapidly in the first century of Fennoscandian boreal forest development. Soil Biol Biochem. 2017; 114: 160–167. https://doi.org/10.1016/j.soilbio.2017.07.017

58. Bellezza S, Paradossi G, De Philippis R, Albertano P. Leptolyngbya strains from Roman hypogea: cytochemical and physico-chemical characterisation of exopolysaccharides. J Appl Phycol. 2003; 15: 193–200. https://doi.org/10.1023/A:1023811311686

59. Sigler W V., Zeyer J. Microbial diversity and activity along the forefields of two receding glaciers. Microb Ecol. 2002; 43: 397–407. https://doi.org/10.1007/s00248-001-0045-5 PMID: 11953808

60. Brankatschk R, Töwe S, Kleineidam K, Schloter M, Zeyer J. Abundances and potential activities of nitrogen cycling microbial communities along a chronosequence of a glacier forefield. ISME J. 2011; 5: 1025–1037. https://doi.org/10.1038/ismej.2010.184 PMID: 21124490

61. Sengupta A, Dick WA. Methanotrophic bacterial diversity in two diverse soils under varying land-use practices as determined by high-throughput sequencing of the pmoA gene. Appl Soil Ecol. Elsevier; 2017; 119: 35–45. https://doi.org/10.1016/j.apsoil.2017.05.031

62. Wang S, Zuo X, Zhao X, Awada T, Luo Y, Li Y, et al. Dominant plant species shape soil bacterial community in semiarid sandy land of northern China. Ecol Evol. 2018; 8: 1693–1704. https://doi.org/10.1002/ece3.3746 PMID: 29435244

63. Sengupta A, Dick WA. Bacterial Community Diversity in Soil Under Two Tillage Practices as Determined by Pyrosequencing. Microb Ecol. 2015; 70: 853–859. https://doi.org/10.1007/s00248-015-0609-4 PMID: 25930203