Soil microbiome in chronosequence of spoil heaps of Kursk Magnetic Anomaly

Ekaterina Ivanova¹,³,⁴, Elizaveta Pershina¹,², Dina Karpova⁵, Olga Rogova³, Evgeny Abakumov², and Evgeny Andronov¹,²,³

¹Federal State Budget Scientific Institution “All-Russian Research Institute of Agricultural Microbiology”, Shosse Podbelskogo, 3, Saint Petersburg, 196608, Russian Federation
²Department of Applied Ecology, Faculty of Biology, Saint Petersburg State University, 16¹ Linia V. O., 29, Saint Petersburg, 199178, Russian Federation
³V. Dokuchaev Soil Institute, Pzhyovskiy Lane, 7 (building 2), Moscow, 119017, Russian Federation
⁴Agrophysical Research Institute, Grazhdanskii Pr., 14, Saint Petersburg, 195220, Russian Federation
⁵Lomonosov Moscow State University, 1-12 Leniniskie Gory, Moscow, 119234, Russian Federation

Address correspondence and requests for materials to Ekaterina Ivanova, ektrnivanova@gmail.com

Abstract

The physicochemical parameters, vegetation and biodiversity of microbiomes inhabiting 10-, 35- and 50-year-old embryonic soils (technozems) formed in the mining areas of the Kursk Magnetic Anomaly (Russia) were analyzed in the current study. Analysis of taxonomic diversity was carried out by 454-pyrosequencing of the V4 variable region of the 16S rRNA gene. All points of the chronosequence were characterized by microbiomes reliably differing in taxonomic composition. The older the dump, the higher the proportion of bacteria from the phyla Actinobacteria (mostly representatives of the Solirubrobacteriaceae family), Chloroflexi and Acidobacteria in the community. In 10-year-old dumps, bacteria from the phyla Proteobacteria, Gemmatimonadetes and Bacteroidetes prevailed. The most pronounced changes in the community structure at the early and middle stages of microbiological succession were demonstrated by bacteria from the genus Delftia, which may indicate their active role in the processes of soil formation in this ecosystem.

Keywords: microbial communities, embryonic soils, 454-pyrosequencing, 16S rRNA, overburden rock dumps.

Introduction

Soil genesis is closely associated with the activity and evolution of its microbiome (Wolitska et al., 2018). Microorganisms play a key role in ensuring the cycles of the basic biogenic elements, in the processes of decomposition of organic material, and in the formation of soil organic matter. Also, soil microorganisms are able to utilize a wide range of chemical compounds, thereby participating in the bioremediation of disturbed areas (Schütz et al., 2010; Dzonek, Wojcieszy and Gzik, 2016; Wolinska et al., 2018). Because of the important ecosystem functions mediated by microorganisms in the soil, recovery of the soil microbial community is a critical step in achieving the goal of soil restoration for its sustainable and beneficial use (Rana, Stahl, Ingram, and Wick, 2007).

The promising model objects are soil chronosequences, where the processes of soil restoration can be observed in development within certain time intervals. A number of such objects are well known, including Antarctic soils (Nemergut et al., 2007; Bradley, Singarayer, and Anesio, 2014), volcanic ash-covered areas (Gomez-Alvarez, King, and Klaus, 2007), post-fire ecosystems (Gobena et al., 2005; Sun et al., 2017), as well as the areas of mining and storage for rock dumps (Spro-
cations (KMA) was a mixture of tertiary Callovian clays (Upper Jurassic) (dominating) and Bat sands. For the analysis, sites were chosen on the dumps of 10 years, 35 years and 50 years of age (CL1-3, respectively, N 52,2592436; E 35,3708321). The soil samples were taken from the upper soil horizon of each of them (at a depth of 0–10 cm). All samples were accompanied by a geobotanical description of the sites, as well as by the main soil physicochemical parameters (Table 2). Sampling was carried out in triplicate, from the middle of the 0–10 cm layer.

DNA extraction was performed using the PowerSoil® DNA Isolation Kit (MO BIO, USA), including the step of mechanical destruction of a soil sample using abrasive materials (Mobio Laboratories, USA). The destruction of the soil sample was carried out on a Precelys 24 homogenizer (Bertin Technologies, France). The purity of isolation and the amount of isolated DNA were checked by electrophoresis in 1 % agarose in 0.5x TAE buffer. The average DNA concentration in the sample was 50 ng/ml.

The purified DNA preparations (10–15 ng) were used as templates in the PCR reaction (temperature profile: 95 °C — 30 s, 50 °C — 30 s, 72 °C — 30 s, 30 cycles in total) using Encyclo polymerase (Eurogen, Russia) and universal primers to the V4 variable region of the 16S rRNA gene: F515 (GTGCCAGCMGCCGCGGTAA) and R806 (GGACTACVSGGGTATCTAAT). The primers included oligonucleotide barcodes for each sample and the service sequences required for 454-pyrosequencing by Roche protocol (Roche, Switzerland). Sample preparation and sequencing were performed on a GS Junior device (Roche, Switzerland) according to the manufacturer’s recommendations.

Downstream processing of the 16S rRNA gene libraries was performed using the QIIME software package (Caporaso et al., 2010). At the first stage, the sequence quality was verified. Sequences smaller than 200 nucleotides in length, with a quality score of less than 25, containing incorrectly read primer and multiplex identifier sequences, as well as extended homopolymeric repeats (more than eight nucleotides) and unidentified nucleotides were excluded from the analysis. Also, all non-bacterial and chimeric sequences were excluded from further analysis. The data were normalized according to the number of sequences in the library of the smallest size. As a result of all of these procedures, 14,312 sequences were selected. After the data normalization procedure, the number of sequences in each library was 4700. Sequences with similarity over 97 % were combined into operational taxonomic units (OTUs) using a de novo algorithm (based on the UCLUST method). From each OTU, one sequence was chosen to compose a set of representatives. The next
Table 1. Vegetation description of the embryozems formed on the overburden dumps of KMA

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Parent material</th>
<th>Plant community composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC_35</td>
<td>Tree tier: <em>Pinus sylvestris</em>, <em>Betula pendula</em>; undergrowth: <em>Pinus sylvestris</em>, <em>Populus tremula</em>; grassy tier — meadow-moatley grass community: <em>synusia of Calamagrostis canescens</em>, <em>Pilosella officinarum</em>, <em>Cirsium vulgare</em>; <em>Querqus robur</em>, <em>Picea abies</em>, <em>Silene pratensis</em>, <em>Fragaria vesca</em>, <em>Linaria vulgaris</em>, <em>Trifolium repens</em>, <em>Achillea millefolium</em>; a significant area was covered with two types of moss.</td>
<td></td>
</tr>
<tr>
<td>CC_50</td>
<td>Plantations of <em>Betula pendula</em> and <em>Populus tremula</em>, the formula of the stand: SB5O (B — birch, O — aspen), a single plant of <em>Pinus sylvestris</em>, underbrush of <em>Querqus robur</em>; <em>Cirsium vulgare</em>, <em>Silene pratensis</em>, <em>Fragaria vesca</em>, <em>Linaria vulgaris</em>, <em>Vicia cracca</em>, <em>Trifolium medium</em>, <em>Achillea millefolium</em>, <em>Calamagrostis canescens</em>.</td>
<td></td>
</tr>
</tbody>
</table>

stage was the classification of representative sequences using the RDP naive Bayesian RNA Classifier program and PyNast algorithm alignment (Caporaso et al., 2010), and a specially designed set of sequences, Greengenes coreset (DeSantis et al., 2006) was used to align the sequences. The aligned sequences were used to construct the distance matrix and phylogenetic tree.

To assess biodiversity and conduct comparative analysis of communities, the parameters of α- and β-diversity were calculated. The α-diversity was assessed using the index of species richness (the number of OTUs in the sample) and the Shannon index. The reliability of differences in α-diversity indices among microbiomes was assessed using the t-test. To assess the β-diversity, the weighted unifrac method was used, which allows estimating the percentage of similarities/differences between all pairs of compared microbiomes (Caporaso et al., 2010). To represent the results of the analysis, the methods of multidimensional statistics (analysis of the main components) using the Emperor program were employed.

To assess the reliability of differences in the representation of individual taxa in the analyzed samples, in addition to the QIIME software, a script was written using the programming language Python. This script performs multiple pairwise tests of contingency tables of OTU frequencies in different collections of samples/replicates. The algorithm dynamically chooses either the G-test or the Fisher exact test depending on the circumstances and applies Bonferroni p-value correction.

**Results and Discussion**

**VEGETATION**

Technozems (Lithosols Technic) of the very initial staged dumps were dominated by grasses, while the sparse forests with dominance of birch and poplar were typical for older stages. During the study, 22 plant species were described (Table 1). The number of species was relatively low in comparison to the previous geobotanical observation of the territory (Stifiev, Golovastikova, and Bessonova, 2011). The reason can be the season when the study was conducted: soil sampling was carried out in September, when the majority of plants are already at the terminal vegetation stage. The major groups are partially similar with species described previously (Stifiev and Golovastikova, 2015). However, only CC_50 samples (50-year dumps) were characterized by 100% plant cover.

**SOIL PHYSICOCHEMICAL CHARACTERISTICS OF THE SAMPLES**

Soils of the studied dumps were presented mainly by weakly developed profiles of Lithosols (embryozems) with well-pronounced accumulation of organic matter in the topsoil (at late stages of succession). No significant alteration of the majority of soil physicochemical properties, as well as the soil structure and texture composition with the age of the dumps, were revealed. The embryozems were generally characterized by an increased content of trace elements, including heavy metals (TM) in comparison with values of MAC, APC (at pH>5.5) and background value (Brindukova, 2010) in the region.

The pH of embryozems was slightly alkaline and fluctuated around 8.0, while the pH of the pristine soil environment was slightly acid. The reason is the composition of the rocks, for instance, a relatively high content of carbonates here.

Increased levels of available potassium can be attributed to the high ability of clays to absorb potassium in a non-exchangeable form (Kovda, 1965). The amount of available phosphorus did not change significantly.

Iron content was high both in the soils on the embryozems and in the background soils, which is associated with the particular features of the KMA region. At
the same time, there was only a weak tendency for the iron content to change with time (Table 2). The lack of correlation of iron content with age is probably due to the low mobility of iron under conditions of natural and slightly alkaline pH (Brindukova, 2010).

Nitrogen and carbon content increased with soil age. Nitrogen was actively consumed by vegetation, and its amount in the soil was related to its content in organic compounds. Moreover, the dumped rocks almost lack nitrogen, and its accumulation in soil occurs due to symbiotic and free-living microorganisms (Nemergut et al., 2007). Therefore, an increase in the nitrogen content in embryozems from the earliest to the latest stages of technogenic landscapes is associated with a similar accumulation of organic carbon. The C:N ratio in the embryozems displayed a low and very low enrichment of humus with nitrogen, which is typical for horizons containing slightly humified plant residues.

According to the Mantel test results, almost all of the main physicochemical parameters caused a significant impact on the structure of microbial community (Table 3).

### ANALYSIS OF PROKARYOTIC COMMUNITY OF THE SOIL CHRONOSEQUENCE

When analyzing the taxonomic structure of the microbiome, 29 bacterial and 2 archaial phyla were found in the soil samples of the dumps. On average, 0.5% of the sequences were not identified to the domain level. Among the bacteria in the community, representatives of phyla Proteobacteria (41.7% on average), Actinobacteria (34%), Chloroflexi (10.4%), Gemmatimonadetes (4.5%), Acidobacteria (2.5%), Bacteroidetes (2%), Planctomycetes (1.1%) (Fig.1a) were noted as dominants; meanwhile other phyla on average did not exceed 1% per library.

When comparing the early and late stages of the investigated microbial succession, the percentage and taxonomic diversity of bacteria from phyla Actinobacteria, Chloroflexi and Acidobacteria increased in the community coincided with increasing age of the dump, while the proportion of such phyla as Proteobacteria, Gemmatimonadetes and Bacteroides decreased. Also, with increasing age of the dump, the number of *Archaea* belonging to the phyla Crenarchaeota and Euryarchaeota decreases almost to zero. The decrease in *Archaea* abundance may be due to the replacement of their ecological niches (in particular, for the representatives of the phylum Crenarchaeota participating in nitrification processes) by bacteria in a more developed type of microbial community (Weidler, Gerbl and Stan-lotter, 2008).

Changes in the taxonomic composition of the prokaryotic community are associated with the transition from an unstable stage of development of an ecosystem characterized by the dominance of copiotrophic microorganisms (for example, bacteria from the genus *Pseudomonas*, bacteria from the family *Enterobacteriaceae*) to the climax type of a community with a predominance of Actinobacteria, the presence and high biodiversity of which indicate an effective process of transformation of soil organic matter (Goldfarb et al., 2011). In the 50-year-old dumps, the proportion of Streptomycetes and Nocardia increases; in particular, a reliable increase in the proportion in the community (*p* < 0.05) was found for families Micromonosporaceae, Nocardioidaceae, Microbacteriaceae and order Solirubrobacterales typical in the soil of areas with a high value of the hydrothermal coefficient. The highest number of nucleotide sequences of the 16S rRNA gene in the analyzed amplicone libraries was shown to belong to Proteobacteria from the Comamonadaceae family. All of them showed the greatest

### Table 2. Physicochemical analysis of samples of the embryozems formed on the overburden dumps of KMA

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CC_10</td>
<td>19.28</td>
<td>6.80</td>
<td>1.15</td>
<td>0.06</td>
<td>5.70</td>
<td>6.32</td>
<td>23.52</td>
<td>15</td>
<td>31</td>
<td>0</td>
<td>57</td>
<td>49</td>
</tr>
<tr>
<td>CC_35</td>
<td>24.99</td>
<td>8.13</td>
<td>1.46</td>
<td>0.05</td>
<td>14.37</td>
<td>1.79</td>
<td>47.96</td>
<td>15</td>
<td>33</td>
<td>14</td>
<td>74</td>
<td>108</td>
</tr>
<tr>
<td>CC_50</td>
<td>22.78</td>
<td>8.12</td>
<td>2.41</td>
<td>0.20</td>
<td>15.95</td>
<td>4.85</td>
<td>53.69</td>
<td>23</td>
<td>37</td>
<td>17</td>
<td>88</td>
<td>106</td>
</tr>
</tbody>
</table>

### Table 3. Mantel R statistics values for the samples of embryozems formed on the overburden dumps of KMA

<table>
<thead>
<tr>
<th>Corg</th>
<th>Ntot</th>
<th>pH</th>
<th>Pav</th>
<th>Kav</th>
<th>CaCO₃</th>
<th>Fe</th>
<th>Cu</th>
<th>Ni</th>
<th>Pb</th>
<th>Zn</th>
<th>Sr</th>
<th>Rb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mantel R statistics</td>
<td></td>
<td>0.38</td>
<td>0.44</td>
<td>0.79</td>
<td>0.76</td>
<td>0.15</td>
<td>0.83</td>
<td>0.83</td>
<td>0.39</td>
<td>0.04</td>
<td>0.89</td>
<td>0.92</td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>0.028</td>
<td>0.010</td>
<td>0.011</td>
<td>0.003</td>
<td>0.349</td>
<td>0.003</td>
<td>0.004</td>
<td>0.024</td>
<td>0.805</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
</tr>
</tbody>
</table>
similarity with the nucleotide sequences of the typical strain of *Delftia tsuruhatensis*. It is known that this microorganism is often associated with contaminated and technogenically disturbed soils, as well as with the microbial community of activated sludge (Hou et al., 2015). Possible participation of these bacteria in soil-forming processes was also confirmed by a peak-like change in their representation in the samples (Fig. 1b). For example, the number of these bacteria started to increase already in the 10-year-old dumps and reached a maximum in the 35-year-old dumps, after which it decreased.

It can be assumed that microbiological succession occurs even 35 years after the formation of the dump. This conclusion is confirmed by the analysis of the alpha and beta diversity of the studied microbiomes (Table 4).

**Table 4. Indicators of alpha diversity of microbiomes inhabiting the embryozems formed on the overburden dumps of KMA. Data are presented as Mean ± SE**

<table>
<thead>
<tr>
<th>Indices</th>
<th>CC_10</th>
<th>CC_35</th>
<th>CC_50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed OTUs*</td>
<td>373.3±139.0</td>
<td>370.3±49.7</td>
<td>466.3±9.6</td>
</tr>
<tr>
<td>PD_Faith**</td>
<td>38.6±10.5</td>
<td>35.6±3.9</td>
<td>39.5±0.6</td>
</tr>
<tr>
<td>Shannon***</td>
<td>6.5±0.9</td>
<td>6.6±0.6</td>
<td>7.6±0.1</td>
</tr>
</tbody>
</table>

* Observed OTUs — total number of operational taxonomic units
** PD_Faith — an indicator of the phylogenetic diversity of Faith
*** Shannon — Shannon Index

Young (10-year-old and 35-year-old) dumps were characterized by similar parameters of alpha diversity (Table 4), indicating an unstable stage of ecosystem
development. The number of species in the community, as well as the Shannon index, increased only in the 50-year-old dumps, which, according to these indirect indicators, can be characterized as the phase transitional to the termination (climax) stage of microbiological succession.

The multidimensional scaling of the distance matrix (Fig. 2) of the samples shows that the structure of microbiomes differed qualitatively and quantitatively in embryozems of different ages.

All the samples formed three separate clusters, according to the age of the dump. It is worth mention that the oldest samples (CC_50) formed the densest cluster, whereas the cluster of the youngest ones was more diffused. This can indirectly indicate more homogenous environment in 50-year old embryozems, which is important for microbiological colonization.

**Conclusions**

Based on the analysis of microbiological diversity, we evidenced that technogenic dumps formed by clay textured overburden rocks in the KMA region could be treated as a favorable niche for the development of microorganisms already at the earliest stages of soil formation. Even in 10 years embryonic soil microbial communities demonstrate a relatively high level of biological diversity. However, the increase in biodiversity indicators and the transition to a stable (climax) type of community were observed only 50 years after the formation of the dumps.

This is characterized by the replacement of copiotrophic forms of microorganisms by oligotrophic ones, among which the main role is played by various types of Actinobacteria.

**Acknowledgements**

The authors are grateful to A. Stefeev, Dr. S.-H. Sciences, Academician of the Academy of Natural Sciences, Professor FSBEI HPE FSBEI HPE Kursk State Agricultural Academy, for assistance in the selection of objects and the selection of samples for research.

**References**


Claassens, S., Rensburg, P. J. J. V. A. N., and Rensburg, L. V. A. N. 2006. Soil microbial community struc-


