First insight into microbial community composition in a phosphogypsum waste heap soil

Sylwia Zielińska[1#✉], Piotr Radkowski[2], Tadeusz Ossowski[3], Agnieszka Ludwig-Gałęzowska[2], Joanna M. Łoś[1] and Marcin Łoś[1]

[1]Department of Molecular Biology, University of Gdansk, Gdańsk, Poland, [2]Center for Medical Genomics – OMICRON, Jagiellonian University Medical College Faculty of Medicine, Kraków, Poland; [3]Department of Analytical Chemistry, University of Gdansk, Gdańsk, Poland

The aim of this study was to investigate the soil microbial communities of a phosphogypsum waste heap. The soil microbial community structures can differ over time, as they are affected by the changing environmental conditions caused by a long-term exposure to different kinds of pollutions, like is the case of soil in the post-production waste area in Wiślinka (in the northern part of Poland) currently undergoing restoration. Our analyses indicated that the most abundant phyla were *Proteobacteria*, *Acidobacteria*, and *Actinobacteria,* and generally such an abundance is common for most of the studied soils. The most dominant class were Alphaproteobacteria, with their participation in 33.46% of the total reads. Among this class, the most numbered order was *Sphingomonadales*, whereas among this order the *Sphingomonadaceae* family was the most abundant one. The *Sphingomonadaceae* family is currently in the center of interest of many researchers, due to the ability of some of its members to utilize a wide range of naturally occurring organic compounds and many types of environmental contaminants. This kind of knowledge about microbial populations can support efforts in bioremediation and can improve monitoring changes in the contaminated environments.

✉e-mail: sylwia.zielinska@biol.ug.edu.pl
#Present address: Department of Bacterial Molecular Genetics, University of Gdansk, Gdańsk, Poland
Abbreviations: NGS, Next Generation Sequencing

INTRODUCTION

Microbial communities in the natural environment are usually very complex and consist of many species. Vast majority of the species present in such communities cannot be propagated in the laboratory in the form of pure culture. Some estimations say that we lack the ability to culture almost 99% of all bacteria using traditional microbiology methods. Even improved methods of culturing soil bacteria allow for culturing only up to 14.1% of all bacterial cells present in a given sample, which leaves the vast majority of cells uncultured (Janssen *et al.*, 2002). These observations facilitated the development of new approaches and methodologies in order to identify the majority of microorganisms. Recent advances in high-throughput sequencing technologies provide a glimpse into to the microbial community structures in a variety of environments, including soil. This method has revealed a high diversity of microbes in different niches and improved our understanding of the new taxa, their biogeographic distribution and the association of specific microbial groups with geochemical factors (Guan *et al.,* 2013) or various kinds of pollutions (Ceja-Navarro *et al.,* 2010; Kravchenko *et al.,* 2010; Azarbad *et al.,* 2015).

Soil is often described as the most challenging of all natural environments, mostly because it is a highly complex environment with the diverse group of microorganisms. The variety of enzymatic inhibitors, like humic acids or heavy metals, often co-extracted with DNA, may interfere with the sample analysis, as they are well known DNA polymerase inhibitors (Daniel, 2005). Moreover, problems also occur because of mineral particles of different size and origin, as well as due to organic compounds that can be present at various stages of decomposition (Daniel, 2005). Each soil environment can be characterized by various factors, for e.g. contamination with pollutants or high influence of anthropogenic factors. Those factors can cause problems not only during the DNA extraction step, but also during further analysis and data interpretation (Felczykowska *et al.,* 2015). Until now, many studies have been conducted in order to determine the influence of different kinds of pollutions, for e.g. metal or diesel contaminations, on the diversity of microbial communities (Ferrera-Rodríguez *et al.,* 2013; Gołębiewski *et al.,* 2014; Gray *et al.,* 2015).

In this study, we present the variability of the soil microbial populations in a post-production waste area, i.e. a phosphogypsum waste heap at Wiślinka, Poland (about 10 km from Gdańsk – part of the Żuławy Gdańskie). Since 1969, a selected area in Wislinka became a landfill for post-production wastes. Phosphorites, which are the main component in the production of fertilizers, were transported to the factory in Gdansk from North Africa. Phosphogypsum was derived as a postproduction waste in the fertilizer production process, and was stored at the landfill in Wiślinka (Skwarzec *et al.,* 2010). The size of the actual landfill (area of the heap) is 26 ha, but together with the protection zone, it amounts to 85 ha (Hupka, 2006). Currently, this area is subjected to reclamation, and was covered with sludge from wastewater treatment plants, which is characterized by the presence of heavy metals, such as copper, zinc, cadmium, lead, chromium and nickel; the area is also covered with vegetation. As phosphogypsum waste heap is a unique place on European scale, in recent years, a number of analyses were conducted in the neighborhood of this landfill in order to determine its impact on the local environment and ecosystem (Skwarzec *et al.,* 2010; Borylko & Skwarzec, 2013; Borylo & Skwarzec, 2014).

The aim of this study was to investigate the soil microbial community structure of the phosphogypsum waste heap with the use of the Next Generation Sequencing (NGS) analyses of the 16S rRNA gene. We expected a complex structure of this particular microbial community, mostly due to the composite structure of the landfill, as well as the postproduction waste and sewage sludge with the cover of vegetation. We wanted to estimate and define a microbial population's structure that would be characteristic for the polluted soil environments, as the long-term exposure to pollutions can lead to many changes and adaptations in the microbial community's structure.

MATERIALS AND METHODS

Study area and sample collection. The soil sample was collected in the north part of Poland (Wiślinka; 54° 19' N, 18° 50' E) at the beginning of September 2015. The company managing the landfill did not consent to soil sampling directly from the landfill area, and thus the sample was collected from the area directly adjacent to the phosphogypsum waste heap. Area surrounding the heap has similar composition as the area of the heap itself. At the time of waste gathering, the area was unprotected and wind and rain relocated phosphogypsum wastes and soil from the strict area of the heap. Even now, there are different reports and expert opinions that strongly suggest that the area of the heap is not protected properly and wastes are being distributed around, not only within the village of Wislinka and surrounding area, but also with the water of the Vistula river directly to the sea. The soil sample was collected from an area equal to one square meter, after ground vegetation was removed, with the use of the Eijkelkamp collection soil rings. Samples from eight single rings were combined and randomly put together, mixed and sieved to form a composite sample. Plant roots were removed from the soil. The depth of the soil layer was above 10 cm. Immediately after collection (within an hour), the soil sample was transported in a sterile container for sieved soil, into a lab facility. Then, for the DNA extraction, the sample was divided into smaller portions and processed according to the DNA extraction kit manufacturer's instructions and for physico-chemical soil analysis.

Physicochemical analysis. The soil physicochemical features: soil dry mass, soil conductivity, pH and nitrogen (NO_3^- and NH_4^+) and phosphate (PO_4^{3-}) content were measured according to the procedure published by Zwolicki and coworkers (2013; 2015). The organic matter content was established according to the protocol published by Myślińska (1998). The total sulphur content was established according to the standard PN-76 B06714 (determination of sulphur content, the bromine method). The concentration of metals (Cu, Ni, Zn, Cr, Cd, and Pb) in the soil sample was determined by using Atomic Absorption Spectrometry in an acetylene-air flame (spectrometer AAnalyst 400, Perkin Elmer). The wavelengths for metals were: Cu 324.8 nm; Ni 232.0 nm; Zn 213.9 nm; Cr 357.87 nm; Cd 228.80 nm, and Pb 283.31 nm. All chemicals used were of an analytical reagent grade. All laboratory ware used in the analyses was previously soaked in a nitric acid 4 mol/l solution for at least 24 h and rinsed with deionised water.

DNA extraction. DNA was extracted from the soil sample with the use of 8 commercially available kits according to the manufacturers' protocols – detailed extraction procedure is available in Zielińska and cowork-ers (2017). To avoid cross contamination of the samples, the process was performed with sterile equipment. The quantity and quality of the extracted DNA were evaluated by using a Nano Drop 1000 spectrophotometer (Nanodrop Technologies). After extraction, the DNA was stored at –20°C until further use.

16S rRNA gene amplification and sequencing. The V3-V4 hypervariable regions of bacterial 16S rRNA gene region were amplified using the following primer set: 341F – CCTACGGGNGGCWGCAG and 785R – GACTACHVGGGTATCTAATCC. The targeted gene region has been shown to be the most appropriate for the Illumina sequencing (Klindworth et al., 2013). Each library (3 separate libraries for each extraction method) was prepared with a two-step PCR protocol based on Illumina's "16S metagenomic library prep guide" (15044223 Rev. B), using Q5® Hot Start High-Fidelity 2X Master Mix (NEBNext – New England BioLabs, PCR under the following conditions: 98°C for 30 sec for initial denaturation of the DNA, followed by 25 cycles of 98°C for 10 sec, 55°C for 30 sec, 72°C for 20 sec and additionally 72°C for 2 min), and the Nextera Index kit (2×250 bp). Paired-end (PE, 2×250 nt) sequencing with a 5% PhiX spike-in was performed with an Illumina MiSeq (MiSeq Reagent kit v2) at Genomed, Warsaw, Poland; following the manufacturer's run protocols (Illumina, Inc., San Diego, CA, USA). For each extraction method, 3 sequencing replicates were performed (Zielińska et al., 2017). The automatic primary analysis and the de-multiplexing of the raw reads were performed with MiSeq, with the use of MiSeq Reporter (MSR) v2.4 (BaseSpace).

Sequencing data analysis and statistical analysis. The samples were processed and analyzed using the Quantitative Insights Into Microbial Ecology (Qiime) pipeline v 1.9.1 software (Caporaso et al., 2011). Low quality PE reads (S. 2010) and adapter sequences (Martin, 2011) were removed from further analysis. Quality-filtered reads were merged based on the overlap of PE read with the use of fastq-joint (Aronesty, 2011). The remaining sequences that did not meet the quality criteria, were removed from further analysis (mean quality >20). Clustering of operational taxonomic units (OTUs) at 97% similarity was performed by using the uclust method, version 1.2.22q (Edgar, 2010). OTUs were assigned to taxa using the GreenGenes release 13.08 as the reference (DeSantis et al., 2006), with taxonomy assignment tool PyNAST (Caporaso et al., 2010). The Biological Observation Matrix (BIOM) table was used as the core data for downstream analyses (McDonald et al., 2012) and vsearch 1.7.0 (VSEARCH GitHub website: https://github.com/torognes/vsearch) as OSS replacement of usearch 6.1. Based on clusters, the diversity indices were estimated, including the Chao1, PD (a quantitative measure of phylogenetic diversity), Shannon, and Simpson indices and also the observed OTUs. Comparison of the microbial community structure was performed with the use of UniFrac (Lozupone & Knight, 2005) and Emperor (Vázquez-Baeza et al., 2013). NGS data are deposited and fully available under study accession number PRJEB12454 in ENA – the European Nucleotide Archive.

For DNA extraction, various commercial kits were used in order to reveal the real structure of the microbial community (Zielińska et al., 2017). Due to such DNA extraction strategy, in further stages we obtained various numbers of the 16S rRNA reads. Those numbers of reads cannot be directly compared to each other in order to reveal the microbial structure, and they cannot

be combined as one sample, as it would generate bias in the data analysis and interpretation. It could lead to overrepresentation of some taxa and underrepresentation of others in the tested microbial population. Therefore, in order to compare and standardize the data and to enable their interpretation, we calculated the mean value for all tested commercial kits with the data presented in percentages.

RESULTS

Physicochemical analysis

Organic matter content of the tested soil sample was at 1.63% and the dry mass at 89%. Conductivity was equal to 92.05 mS/cm. The values of soil PO_4^{3-}, NO_3^- and NO_3^- ion content (mg/kg soil dry mass) were 206, 0.04 and 0.002, respectively. The content of sulphur was at 0.03%. The soil pH was 7.06. The concentration of the following metals: Cu, Ni, Zn, Cr, Cd, and Pb (mg/kg) in the soil sample was 0.111, 0.294, 0.528, 0.276, 0.045, and 0.669, respectively.

General description of the sequencing results

On average, we obtained 160 443 raw reads and 64 150 good quality rRNA gene sequences (V3-V4 region) per each sequenced sample. We were able to classify all of the obtained sequences at the phylum level. When comparing all data together, the value of the Shannon index was 8.00 ± 1.900, the value of the Simpson index was 0.963 ± 0.059, while the value of the Chao1 index was 2168.726 ± 1008.150 and PD was 59.480 ± 23.240. The diversity indices represent a randomly selected subset for a sample normalized to 26,090 sequences. The number of observed OTUs was $1\,631.930 \pm 732.149$. Detailed taxonomic analyses at different ranks are available in the Supplementary Data, presented as a sunburst chart for average values calculated for data obtained for all of the tested kits (Supplementary Fig. 1 at www.actabp.pl) and also as a table (Supplementary Table 1 at www.actabp. pl). Detailed data for single DNA extraction with the use of different extraction kits are in Zielińska and coworkers (2017).

Microbial community composition

The analysis of microbial community found in the phosphogypsum waste soil sample demonstrated that 99.99% of the total reads were represented by *Bacteria* and 0.01% by *Archaea* (Supplementary Fig. 1 and Supplementary Table 1 at www.actabp.pl). Taxonomy-based analysis showed that the soil microbial community consisted of 33 phyla (Supplementary Fig. 1 and Supplementary Table 1 at www.actabp.pl). The most abundant
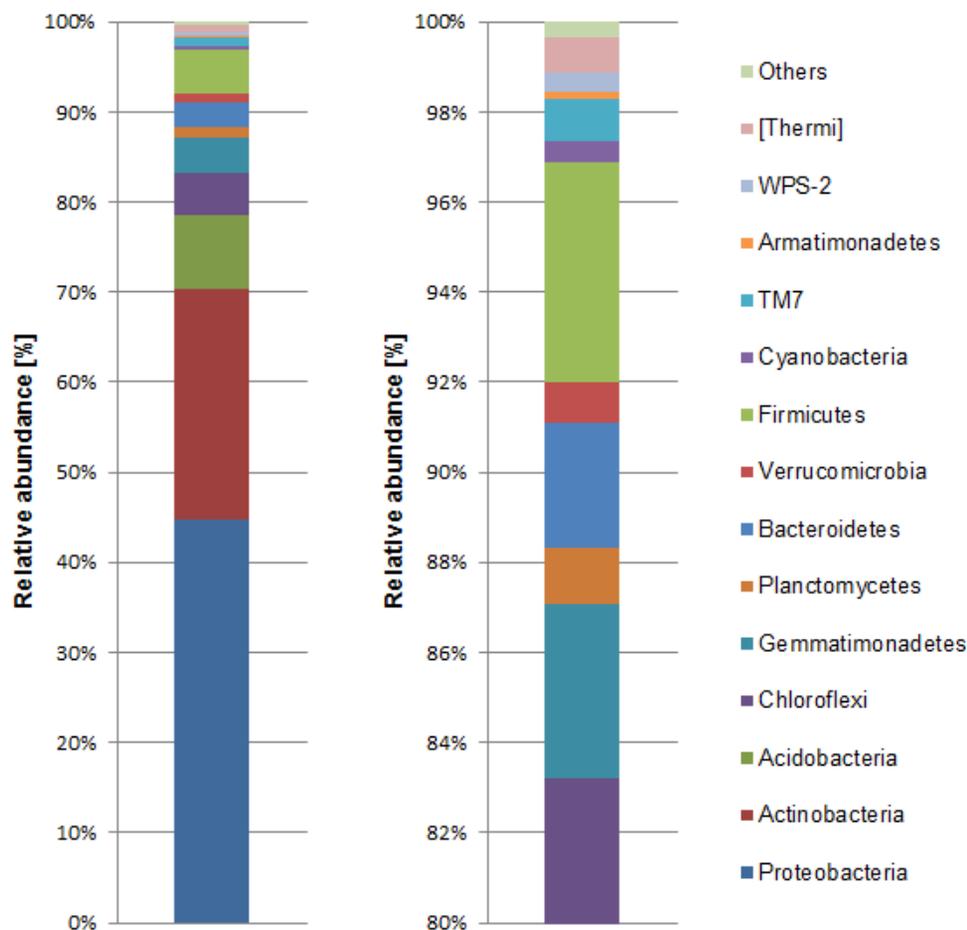

Figure 1. Abundance of microbial 16S rRNA sequences at the phylum level.
Analysis of the mean values of microbial community structure. "Other" describes: TM6, *Nitrospirae*, OD1, *Chlorobi*, *Crenarchaeota*, Elusimicrobia, *Planctomycetes*, *Fibrobacteres*, FBP, MVP-21, *Tenericutes*, WS2, AD3, *Chlamydiae*, BRC1, OP11, *Spirochaetes*, *Fusobacteria*, Euryarchaeota, FCPU426. Left panel shows total microbial community (0–100%) in soil sample and right panel presents the clearly less abundant phyla (84–100%).

phyla were *Proteobacteria*, *Acidobacteria*, and *Actinobacteria* (Fig. 1, Supplementary Fig. 1 and Supplementary Table 1 at www.actabp.pl). Those phyla jointly accounted for more than 78% of the total microbial sequences, followed by *Firmicutes* (4.86%), *Chloroflexi* (4.68%) and *Gemmatimonadetes* (3.8%). Jointly, 12 phyla were responsible for more than 99% of the total microbial population.

In our taxonomy analysis of the soil microbial community, we were able to identify 96 classes, 138 orders, 183 families and 264 genera. 9 classes among bacteria remained unclassified. *Alphaproteobacteria* were the most dominant class, with their participation in 33.46% of the total reads. The second most abundant class were *Actinobacteria*, with their share of 16.04% in the total reads in the microbial population. Additional 7 classes: *Gammaproteobacteria*, *Thermoleophilia*, *Acidobacteriia*, *Bacilli*, *Betaproteobacteria*, *Acidimicrobiia*, and Ellin6529 separately contributed to more than 2.5% of the total share, and together those 9 classes accounted for more than 78% of the total reads (Supplementary Fig. 2 at www.actabp.pl). Among the most dominant class of *Alphaproteobacteria*, the most numerous order was *Sphingomonadales*, with contribution of 19.45% in the total microbial population. The second largest order was *Rhizobiales*, which constituted 8.82%. Among the *Sphingomonadales* order, the *Sphingomonadaceae* family was the most abundant, with the contribution of 18.89%. On the other hand, among the *Sphingomonadaceae* family, *Kaistobacter* was the most abounded genus with the share of 16.51%, followed by *Sphingomonas* with the total share of 1.65%.

DISCUSSION

Due to changing environmental conditions, the soil microbial communities are expected to contain various levels of microbial species, and in consequence constitute a valuable source of information about the functioning of a given ecosystem. Specifically, it is believed that understanding of the soil microbial communities' structures can be crucial in predicting the response of the soil environment to the climate changes and ecological disasters related to the changes in the pollution levels, or even in the reclamation processes. Therefore, many metagenomics studies focus on determining the microbial community structures and their relationship to the changing soil properties (Azarbad *et al.*, 2015; Campbell *et al.*, 2010; Ramirez *et al.*, 2010; Kim *et al.*, 2014).

Chu and coworkers (2010) had demonstrated, with an example of the Arctic soil bacterial communities, that those bacterial communities are characterized with similar levels of variability, richness and phylogenetic diversity as soil samples from a wide range of lower latitudes. This suggests a common diversity structure within the soil bacterial communities around the globe. It also suggested that the Arctic soil bacterial communities are strongly influenced by local environmental factors, especially by the soil acidity. This can lead to a conclusion that the established biome definitions are not useful for predicting variability in the soil bacterial communities across the globe, as it is in the case of well-established latitudinal gradients in animal and plant diversity. Azarbad and coworkers (2015) had shown, based on the 16S rRNA gene sequencing and GeoChip data, that the soil microbial communities adapted to the elevated metal concentrations in the polluted soils. They presented data that the relative abundance of particular groups of microorganism was changed, although the whole composi-

tion of the microbial communities, at the high taxonomic ranks, is common for most of the studied soil samples.

In this study, we have observed a high number of good quality reads, and generally, 15 000–100 000 reads per sample are sufficient for classification, as described by Illumina 16S Metagenomic Sequencing Protocol. Moreover, in this analysis we were able to classify all of the obtained sequences at a phylum level. A relatively high number of OTUs, as well as high values of microbial diversity indexes (Shannon`s and Simpson`s indexes), suggest a high number of species in the tested soil sample. Moreover, the number of OTUs observed in the sample indicates that the microbial population is highly complex, and together with rarefication trends analysis, indicates that sampling of this microbial community is close to being complete. Still, Azarbad and coworkers (2015) had implied, based on their studies where they sequenced about 20 000 amplicons from a metal polluted soil and with over 5 000 different OTUs observed, that this may not be enough for a quantitative representation of a highly diverse microbial community. However, based on the high values of microbial diversity indexes and the high number of OTUs, we suggest that the specific nature of the postproduction waste heap did not reduce the microbial diversity. Additionally, it did not considerably affect the microbial structure at the high taxonomic ranks, when compared to the soil samples of different origin.

The most abundant phyla in the tested sample are: *Proteobacteria*, *Actinobacteria*, *Acidobacteria* and *Chloroflexi*. Generally, such an abundance and composition of share in a microbial community is common for most of the studied soil samples of different origin (Neufeld & Mohn, 2005; Janssen, 2006; Xu *et al.*, 2014). Mostly, minor groups of the microbial community structure were similar to those previously described (Ramirez *et al.*, 2010; Kim *et al.*, 2014; Chu *et al.*, 2010). Furthermore, high abundance of *Alphaproteobacteria* is often associated with an elevated concentration of mobile metals (Reith *et al.*, 2012), especially the presence of *Sphingomonas*, which have a broad physiological tolerance to metals, resulting in a specialized community structure. Additionally, high abundance of the *Alphaproteobacteria* class is linked to a low level of the organic matter present. Low content of C and N in the soil favours genera that are phototropic and/or are capable of nitrogen fixation which makes them ideal inhabitants for a soil sample with a low organic matter content (Reith *et al.*, 2012), as in this case. For this particular soil sample, the microbial population structure is characteristic for polluted soil environments, although the concentration of mobile metals in this soil sample is relatively low, whereas the concentration of phosphate ions is quite high. Nonetheless, the ratio of mobile metals is quite high when compared to the content of organic matter, in the tested sample.

Organisms related to the *Kaistobacter* genus are mainly represented by the *Sphingomonadales* group, and no information is available on their possible function in the soil ecosystems (Ceja-Navarro *et al.*, 2010). However, members of the *Kaistobacter* genus have been detected in methane enriched (Kravchenko *et al.*, 2010), diesel contaminated arctic soils (Ferrera-Rodríguez *et al.*, 2013) and soils containing isoprene (Gray *et al.*, 2015). Moreover, the entire *Sphingomonadaceae* family, due to the ability of some of their species to utilize a wide range of naturally occurring organic compounds, as well as many types of refractory environmental contaminants, is in the center of interest of many scientists, in the context of biore-

mediation of environmental contaminants (Balkwill *et al.*, 2006).

Likewise, an increased abundance of the *Chloroflexi* and *Gemmatimonadetes* phyla (4.68% and 3.84% in the tested soil sample, respectively), is often found at the metal-polluted sites (Azarbad *et al.*, 2015). Unfortunately, little is known about those phyla, especially about *Gemmatimonadetes*, as to their ecology and metabolism. However, the presence and increased abundance of those groups can suggest that they play an important role in the contaminated soils. As those bacteria can be highly adapted to extreme environments, their presence may also provide a stabilizing element in the microbial population (Azarbad *et al.*, 2015). Elevated abundance of the *Firmicutes* phyla, also observed in this study, is frequently found in the contaminated soils, especially in the case of the *Bacilli* class (Ellis *et al.*, 2003; Dell'Amico *et al.*, 2008; Berg *et al.*, 2012).

As microorganisms can play an extremely important role in the decomposition of dead organic matter, any adverse effects in soil parameters, such as pollution, can potentially lead to distortions of the nutrient cycle. This may have an impact on the functioning of the whole ecosystem (Stefanowicz *et al.*, 2009). Therefore, monitoring of the changes in the microbial populations may be an indicator of fluctuations in the natural environment, caused for e.g. by pollutants or environmental disasters. This can be very useful especially when one considers monitoring or even dealing with ecological changes after natural disasters, or at the post-production areas and undergoing restoration. This report gives the first insight into microbial community structure of a phosphogypsum waste heap soil and can be exploit to better understanding which group of microorganisms can be involved in dealing with various contaminations and how the microbial communities deal with those contaminations (Handelsman and *et al.*, Committee on Metagenomics: Challenges and Functional Applications 2007), and can support our efforts in bioremediation of environmental contaminations.

Acknowledgements

REFERENCES

Andrews S (2010) FastQC a Quality Control Tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Aronesty E (2011) Command-line tools for processing biological sequencing data, ea-utils. http://code.google.com/p/ea-utils: Expression Analysis, Durham, NC

Azarbad H, Niklińska M, Laskowski R, van Straalen NM, van Gestel CAM, Zhou J He Z, Wen C, Röling WF (2015) Microbial community composition and functions are resilient to metal pollution along two forest soil gradients. *FEMS Microbiol Ecol* 91: 1–11. doi: 10.1093/femsec/fiu003

Balkwill DL, Fredrickson JK, Romine MF (2006) Sphingomonas and related genera. The prokaryotes, a handbook of the biology of bacteria. Vol 7: *Proteobacteria: Delta and Epsilon Subclasses. Deeply Rooting Bacteria*. Dworkin M, ed. Singapore.

Berg J, Brandt KK, Al-Soud WA, Holm PE, Hansen LH, Sørensen SJ, Nybroea O (2012) Selection for Cu-tolerant bacterial communities with altered composition, but unaltered richness, *via* long-term cu exposure. *Appl Environ Microbiol* 78: 7438–7446. doi: 10.1128/AEM.01071-12

Boryłko A, Skwarzec B (2013) Biogeochemistry of uranium in the southern Baltic ecosystem. *J Ecol Protec Coastline* 92: 53–69

Boryło A, Skwarzec B (2014) Activity disequilibrium between 234U and 238U isotopes in natural environment. *J Radioanal Nuc Chem* 300: 719–727

Campbell BJ, Polson SW, Hanson TE, Mack MC, Schuur EAG (2010) The effect of nutrient deposition on bacterial communities in Arctic tundra soil. *Environ Microbiol* 12: 1842–1854. doi: 10.1111/j.1462-2920.2010.02189.x

Caporaso JG, Bittinger K, Bushman F, DeSantis T, Andersen G, Knight R (2010) PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26: 266–267. doi: 10.1093/bioinformatics/btp636

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2011) QIIME allows analysis of high-throughput community sequencing data. 7: 335–336. doi: 10.1038/nmeth.f.303

Ceja-Navarro JA, Rivera-Orduña FN, Patiño-Zúñiga L, Vila-Sanjurjo A, Crossa J, Govaerts B, Dendooven L (2010) Phylogenetic and multivariate analyses to determine the effects of different tillage and residue management practices on soil bacterial communities. *Appl Environ Microbiol* 76: 3685–3691. doi: 10.1128/AEM.02726-09

Chu H, Fierer N, Lauber CL, Caporaso JG, Knight R, Grogan P (2010) Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environ Microbiol* 12: 2998–3006. doi: 10.1111/j.1462-2920.2010.02277.x

Daniel R (2005) The metagenomics of soil. *Nat Rev Microbiol* 3: 470–478. doi: 10.1038/nrmicro1160

Dell'Amico E, Mazzocchi M, Cavalca L, Allievi L, Andreoni V (2008) Assessment of bacterial community structure in a long-term copper-polluted ex-vineyard soil. *Microbiol Res* 163: 671–683. doi: 10.1016/j.micres.2006.09.003

DeSantis T, Hugenholtz, P, Larsen N, Rojas M, Brodie E, Keller K, *et al.*, (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72: 5069–5072. doi: 10.1128/AEM.03006-05

Edgar R (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460–2461

Ellis R, Morgan P, Weightman A, Fry JAEM (2003) Cultivation-dependent and -independent approaches for determining bacterial diversity in heavy-metal-contaminated soil. *Appl Environ Microbiol* 69: 3223–3320

Felczykowska A, Krajewska A, Zielińska S, Łos JM (2015) Sampling, metadata and DNA extraction – Important steps in metagenomic studies. *Acta Biochim Polon* 62: 151–160. http://dx.doi.org/10.18388/abp.2014_916.

Ferrera-Rodríguez O, Greer CW, Juck D, Consaul LL, Martínez-Romero E, Whyte LG (2013) Hydrocarbon-degrading potential of microbial communities from Arctic plants. *J Appl Microbiol* 114: 71–83. doi: 10.1111/jam.12020

Gołębiewski M, Deja-Sikora E, Cichosz M, Tretyn A, Wróbel B (2014) 16S rDNA pyrosequencing analysis of bacterial community in heavy metals polluted soils. *Microbial Ecol* 67: 635–647. doi: 10.1007/s00248-013-0344-7. doi: 10.1007/s00248-013-0344-7

Gray CM, Helmig D, Fierer N (2015) Bacteria and fungi associated with isoprene consumption in soil. *Elementa: Science of the Anthropocene* 3: 53. doi: 10.12953/journal.elementa.000053

Guan X, Wang J, Zhao H, Wang J, Luo X, Liu F, Zhao F (2013) Soil bacterial communities shaped by geochemical factors and land use in a less-explored area, Tibetan Plateau. *BMC Genomics* 14: 820. doi: 10.1186/1471-2164-14-820. doi: 10.1186/1471-2164-14-820

Committee on Metagenomics: Challenges and Functional Applications (2007) *The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet*. Washington, DC: The National Academies Press

Hupka J (2006) The impact of the phosphogypsum waste dump on the environment and eutrophication of the Gdańsk Gulf. Gdańsk (in Polish)

Janssen P (2006) Identyfying the Dominant Soil Bacterial Taxa in Libraries of 16S rRNA and 16S rRNA Genes. *Appl Environ Microb* 1719–1728. doi: 10.1128/AEM.68.5.2391

Janssen PH, Yates PS, Grinton BE, Taylor PM, Sait M (2002) Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions acidobacteria, actinobacteria. 68: 2391–2396

Kim HM, Jung JY, Yergeau E, Hwang CY, Hinzman L, Nam S, Hong SG, Kim OS, Chun J, Lee YK (2014) Bacterial community structure and soil properties of a subarctic tundra soil in Council, Alaska. *FEMS Microbiol Ecol* 89: 465–475. doi: 10.1111/1574-6941.12362

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* **41**: 1–11. doi: 10.1093/nar/gks808

Kravchenko IK, Kizilova AK, Bykova SA, Men'ko EV, Gal'chenko VF (2010) Molecular analysis of high-affinity methane-oxidizing enrichment cultures isolated from a forest biocenosis and agrocenoses. *Microbiology* **79**: 106–114. doi: 10.1134/S0026261710010145

Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microb* **71**: 8228–8235

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**: 10–12

McDonald D, Clemente J, Kuczynski J, Rideout J, Stombaugh J, Wendel D, Wilke A, Huse S, Hufnagle J, Meyer F, Knight R, Caporaso JG (2012) The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *GigaSci* **1**: 7. doi: 10.1186/2047-217X-1-7

Myślińska E ed (1998) Laboratory soil research. PWN, Warszawa (in Polish)

Neufeld JD, Mohn WW (2005) Unexpectedly high bacterial diversity in arctic tundra relative to boreal forest soils, revealed by serial analysis of ribosomal sequence tags unexpectedly high bacterial diversity in arctic tundra relative to boreal forest soils, revealed by serial Ana. *Appl Environ Microbiol* **71**: 5710–5718. doi: 10.1128/AEM.71.10.5710

Ramirez KS, Lauber CL, Knight R, Bradford MA, Fierer N (2010) Consistent effects of nitrogen fertilization on soil bacterial communities in contrasting systems. *Ecology* **91**: 3414–3463. http://www.ncbi.nlm.nih.gov/pubmed/21302816. doi: doi:10.1890/10-0426.1

Reith F, Brugger J, Zammit CM, Gregg AL, Goldfarb KC, Andersen GL, DeSantis TZ, Piceno YM, Brodie EL, Lu Z, He Z, Zhou J, Wakelin SA (2012) Influence of geogenic factors on microbial communities in metallogenic Australian soils. *The ISME J* **6**: 2107–2118. http://dx.doi.org/10.1038/ismej.201248

Skwarzec B, Boryło A, Kosinska A, Radzajewska S. (2010) Polonium (Po-210) and uranium (U-234, U-238) in water, phosphogypsum and their bioaccumulation in plants around phosphogypsum waste heap at Wislinka (northern Poland). *Nukleonika* **55**: 187–193 (in Polish)

Stefanowicz AM, Niklińska M, Laskowski R (2009) Pollution-induced tolerance of soil bacterial communities in meadow and forest ecosystems polluted with heavy metals. *Eur J Soil Biol* **45**: 363–369. doi: 10.1016/j.ejsobi.2009.05.005

Vázquez-Baeza Y, Pirrung M, Gonzalez A, Knight R (2009) EMPeror: a tool for visualizing high-throughput microbial community data. *Gigascience* **26**: 2–16. doi: 10.1186/2047-217X-2-16

Xu Z, Hansen MA, Hansen LH, Jacquiod S, Sørensen SJ (2014) Bioinformatic approaches reveal metagenomic characterization of soil microbial community. *PLoS ONE* **9**. doi: 10.1371/journal.pone.0093445

Zielińska S, Radkowski P, Blendowska A, Ludwig-Gałęzowska A, Łoś JM, Łoś M (2017) The choice of the DNA extraction method may influence the outcome of the soil microbial community structure analysis. *Microbiol Open* e00453. doi: https://doi.org/10.1002/mbo3.453

Zwolicki A, Barcikowski M, Barcikowski A, Cymerski M, Stempniewicz L, Convey P (2015) Seabird colony effects on soil properties and vegetation zonation patterns on King George Island, Maritime Antarctic. *Polar Biology* **38**: 1645–1655. doi: 10.1007/s00300-015-1730-z