Microbial communities in Pb contaminated boreal forest soil

Nan Hui

Department of Environmental Sciences
Faculty of Biological and Environmental Sciences
University of Helsinki

Academic Dissertation in Environmental Ecology
To be presented, with the permission of the Faculty of Biological and Environmental Sciences, University of Helsinki, for public criticism in the Auditorium of Lahti Science and Business Park, Niemenkatu 73, Lahti on October 29th 2010, at 12 noon
Supervisors
Professor Martin Romantschuk
Department of Environmental Sciences
Faculty of Biological and Environmental Sciences
University of Helsinki
Lahti, Finland

Professor Rauni Strömmer
Department of Environmental Sciences
Faculty of Biological and Environmental Sciences
University of Helsinki
Lahti, Finland

Reviewers
Senior Researcher Jaak Truu
Institute of Molecular and Cell Biology
University of Tartu
Tartu, Estonia

Professor Marko Virta
Department of Food and Environmental Sciences
University of Helsinki
Helsinki, Finland

Opponent
Associate professor Håkan Wallander
Department of Ecology
University of Lund
Lund, Sweden

Custos
Professor Timo Kairesalo
Department of Environmental Sciences
Faculty of Biological and Environmental Sciences
University of Helsinki
Lahti, Finland

ISBN 978-952-10-6559-0 (Paperback)
ISBN 978-952-10-6560-6 (PDF)
ISSN 1799-0580

YLIOPISTOPAINO
HELSINKI 2010
Contents

Abstract 3
LIST OF ORGINAL PAPERS 5
THE AUTHOR’S CONTRIBUTION 5
ABBREVIATIONS 6
1 INTRODUCTION 7
   1.1 Boreal forest soil as a habitat for microbes 7
   1.2 Microbial diversity in boreal forest soils 7
       1.2.1 Bacteria 7
       1.2.2 Fungi 9
       1.2.3 Diversity and function 12
   1.3 Sequencing methods in microbial diversity studies 13
       1.3.1 rRNA genes 13
       1.3.2 Sanger DNA sequencing 14
       1.3.3 Cloning and sequencing 14
       1.3.4 Pyrosequencing 15
       1.3.5 Diversity estimates 15
       1.3.6 Other DNA methods 17
   1.4 Heavy metal contamination of soils 17
       1.4.1 Lead and lead pollution 17
       1.4.2 Shooting ranges in Finland 18
2 AIMS OF THE STUDY 19
3 MATERIALS AND METHODS 20
   3.1 Hälvälä shooting range 20
   3.2 Sampling and methods 20
   3.3 A new program for OTU assignment 23
4 RESULTS AND DISCUSSION 26
   4.1 Microbial communities in Pb contaminated soil 26
   4.2 Ectomycorrhizal fungi 26
   4.3 Community composition of EcM fungi 29
   4.4 Diversity of bacteria and bioavailability of lead 30
   4.5 Lead tolerance of bacteria 31
   4.6 Representativeness of sampling and analysis techniques 32
   4.7 454 pyrosequencing OTU assignment 33
   4.8 The new program for analyzing pyrosequencing data 33
5 CONCLUSIONS 37
6 ACKNOWLEDGEMENTS 38
7 REFERENCES 39
Abstract

Lead contamination in the environment is of particular concern, as it is a known toxin. Until recently, however, much less attention has been given to the local contamination caused by activities at shooting ranges compared to large-scale industrial contamination. In Finland, more than 500 tons of Pb is produced each year for shotgun ammunition. The contaminant threatens various organisms, ground water and the health of human populations. However, the forest at shooting ranges usually shows no visible sign of stress compared to nearby clean environments. The aboveground biota normally reflects the belowground ecosystem. Thus, the soil microbial communities appear to bear strong resistance to contamination, despite the influence of lead.

The studies forming this thesis investigated a shooting range site at Hälvälä in Southern Finland, which is heavily contaminated by lead pellets. Previously it was experimentally shown that the growth of grasses and degradation of litter are retarded. Measurements of acute toxicity of the contaminated soil or soil extracts gave conflicting results, as enchytraeid worms used as toxicity reporters were strongly affected, while reporter bacteria showed no or very minor decreases in viability. Measurements using sensitive inducible luminescent reporter bacteria suggested that the bioavailability of lead in the soil is indeed low, and this notion was supported by the very low water extractability of the lead. Nevertheless, the frequency of lead-resistant cultivable bacteria was elevated based on the isolation of cultivable strains.

The bacterial and fungal diversity in heavily lead contaminated shooting sectors were compared with those of pristine sections of the shooting range area. The bacterial 16S rRNA gene and fungal ITS rRNA gene were amplified, cloned and sequenced using total DNA extracted from the soil humus layer as the template. Altogether, 917 sequenced bacterial clones and 649 sequenced fungal clones revealed a high soil microbial diversity. No effect of lead contamination was found on bacterial richness or diversity, while fungal richness and diversity significantly differed between lead contaminated and clean control areas. However, even in the case of fungi, genera that were deemed sensitive were not totally absent from the contaminated area: only their relative frequency was significantly reduced. Some operational taxonomic units (OTUs) assigned to Basidiomycota were clearly affected, and were much rarer in the lead contaminated areas.

The studies of this thesis surveyed EcM sporocarps, analyzed morphotyped EcM root tips by direct sequencing, and 454-pyrosequenced fungal communities in in-growth bags. A total of 32 EcM fungi that formed conspicuous sporocarps, 27 EcM fungal OTUs from 294 root tips, and 116 EcM fungal OTUs from a total of 8 194 ITS2 454 sequences were recorded. The ordination analyses by non-parametric multidimensional scaling (NMS) indicated that Pb enrichment induced a shift in the EcM community composition. This was visible as indicative trends in the sporocarp and root tip datasets, but explicitly clear in the communities observed in the in-growth bags. The compositional shift in the EcM community was mainly attributable to an increase in the frequencies of OTUs assigned to the genus *Thelephora*, and to a decrease in the OTUs assigned to *Pseudotomentella, Suillus* and *Tylospora* in Pb-contaminated areas when compared to the control. The enrichment of *Thelephora* in contaminated areas was also observed when
examing the total fungal communities in soil using DNA cloning and sequencing technology. While the compositional shifts are clear, their functional consequences for the dominant trees or soil ecosystem remain undetermined. The results indicate that at the Hälvälä shooting range, lead influences the fungal communities but not the bacterial communities. The forest ecosystem shows apparent functional redundancy, since no significant effects were seen on forest trees.

Recently, by means of 454 pyrosequencing, the amount of sequences in a single analysis run can be up to one million. It has been applied in microbial ecology studies to characterize microbial communities. The handling of sequence data with traditional programs is becoming difficult and exceedingly time consuming, and novel tools are needed to handle the vast amounts of data being generated. The field of microbial ecology has recently benefited from the availability of a number of tools for describing and comparing microbial communities using robust statistical methods. However, although these programs provide methods for rapid calculation, it has become necessary to make them more amenable to larger datasets and numbers of samples from pyrosequencing. As part of this thesis, a new program was developed, MuSSA (Multi-Sample Sequence Analyser), to handle sequence data from novel high-throughput sequencing approaches in microbial community analyses. The greatest advantage of the program is that large volumes of sequence data can be manipulated, and general OTU series with a frequency value can be calculated among a large number of samples.
LIST OF ORGINAL PAPERS

This thesis is based on the following papers and a computer program*, which are referred to in the text by their Roman numerals:


II. Hui N., Liu X., Kurola J., Mikola J. and Romantschuk M. Lead (Pb) contamination alters richness and diversity of the fungal, but not the bacterial community in pine forest soil. Submitted manuscript.


THE AUTHOR'S CONTRIBUTION

I. Nan Hui investigated the bacterial community by DNA sequencing, was involved in bacterial culturing experiments and carried out part of the field work. He contributed to the experimental planning, data interpretation and writing.

II. Nan Hui performed the DNA cloning and sequencing work, analyzed bacterial and fungal communities and calculated diversity indices. He planned and wrote the manuscript.

III. Nan Hui characterized EcM fungal communities of pine root tips and in-growth bags by direct sequencing and 454 pyrosequencing. He planned and wrote the manuscript.

*Nan Hui developed the computer program Multi-Sample Sequence Analyzer (MuSSA) with the assistance of Xinyu Tan under the supervision of Martin Romantschuk.
ABBREVIATIONS

ANOVA = Analysis of variance

DNA = Deoxyribonucleic acid

ITS = Internal transcribed spacer. A non-coding sequence between coding ribosomal RNA subunits

HGT = Horizontal gene transfer

OTU = Operational taxonomic unit

PCR = Polymerase chain reaction

PLFA = Phospholipid fatty acid analysis

RNA = Ribonucleic acid

rRNA gene = Ribosomal RNA gene

Sp = Species
1 INTRODUCTION

1.1 Boreal forest soil as a habitat for microbes

Boreal forest soil consists of various components of mineral particles, organic residues and organic matter in various stages of decay. In pore spaces there are water, dissolved minerals, and gases like CO$_2$ and O$_2$. These physical and chemical conditions make the soil an excellent habitat for a complex microbial community. Podzolic soils, which are typical in the boreal zone, have distinctive layers with different biological and chemical properties. Litter from above ground parts of plants accumulate on the soil surface and a major proportion of living roots can be found in the upper 10 cm of soil (Pietikäinen et al. 1999). The humus horizon, below the litter layer, is rich in organic matter and high in microbial biomass. On top of the mineral soil there is the nutrient-poor eluvial horizon beneath which illuvial horizon and parental soil are found. The natural soil environment changes constantly seasonally and with the action of plant roots and soil organisms. The spatially and temporally heterogeneous mixture of various microhabitats harbour a diverse community of soil organisms.

1.2 Microbial diversity in boreal forest soils

Determining the true diversity of microbes has for a long time been seen as a challenge (Schleifer 2004, Ward 2002). There is a huge abundance of microbes on Earth, and the microbes constitute the vast majority of the diversity of life. Although the total number of microbial species has never been counted, estimates suggest anything up to 4000 bacterial species per gram soil, of which at least half and perhaps as many as 95% are as yet uncultivable (Fitter et al. 2005).

1.2.1 Bacteria

In soil, bacteria are able to perform an extremely wide range of chemical transformations, including the degradation of organic matter, disease suppression, disease causation, and nutrient transformations (Kiikkila et al. 2006, Benitez M.S. and Gardener B.B.M. 2009, Astrom et al. 2006, Raynaud et al. 2006). The species concept for prokaryotes is obscure, as it is more typological and less evolutionary than that for plants and animals, and moreover it is much broader and more inclusive. The species in prokaryotes correspond roughly to the genera of eukaryotic organisms, and prokaryotic genotypes correspond to eukaryotic species (Torsvik and Øvreås 2006). Although bacterial do not have a clear species concept as an evolving entity, functional species definitions do exist. Roselló-Mora and Amann (2001) defined bacterial phylo-phenetic species as "monophyletic and genomically coherent cluster of individual organisms that show a high degree of overall similarity in many independent characteristics, and is diagnosable by a discriminative phenotypic property." Recently the degree of DNA: DNA reassociation between strains is utilized to define and distinguish prokaryotic species. Two bacterial organisms have been considered to belong to the same species if their genomes show greater than 70% cross-hybridization (Torsvik and Øvreås 2006).
The two bacterial organisms often exhibit more than 97% of 16S ribosomal RNA (rRNA) gene sequence similarity. Consequently, the 97% 16s rRNA similarity has been proposed to define a bacterial species delimiter (Torsvik and Øvreås 2006). However, organisms with high 16s rRNA relatedness may show considerable genomic differences. The 97% similarity of the 16s rRNA gene sequence has been questioned to define species and regarded as an arbitrary. At present, instead of using the species as a unit, the term operational taxonomic unit (OTU) is then used to indicate those groups that cluster together above certain similarity (often 95-99%) threshold.

In soil, bacteria have traditionally been divided into autochthonous and zymogenous types. Autochthonous microorganisms inhabit the soil at relatively constant levels and many of these are oligotrophic, which means that they grow and metabolize with limited resources at a relatively low growth rate. This may explain why soils support larger numbers and a higher diversity of bacteria than other habitats that are richer in nutrients (van Elsas et al. 2006). On the other hand, zymogenous microorganisms can only grow under the conditions of abundant resources and are usually copiotrophic, meaning that they utilize easily oxidizable organic material. In terms of trophic strategy, K-strategists and r-strategists are ecologically defined and equated to reflect the autochthonous and zymogenous microorganisms. Researchers have studied the kinetics of K- and r-strategists in soil by measuring the response of microbial respiration to glucose amendment, and have assumed that many bacteria in soil can shift from r-type to K-type behaviour (Stenström et al. 1998). These observations indicate that one bacterium can alternate between these two strategies, depending on the available nutrients. In fact, all bacteria, even those considered as copiotrophs, may in practice live with both r- and K-strategies during their cycles (Stenström et al. 2001). In soil, insufficient nutrient conditions may trigger bacteria to shift from an r-strategy to a K-strategy. Although the abundance of such bacteria may dramatically decline, they tend to live with a low growth rate instead of dying away. In this way, the bacterial diversity stays relatively stable and the actual variation in species richness may remain more limited than any measurement would indicate.

Bacteria, probably more efficiently than any other type of organism, may rapidly adapt fluctuating environmental conditions in terms of their genetic and physiological properties (van Elsas et al. 2006). Genetic adaptation is often related to horizontal gene transfer (HGT), meaning that genes are transferred from a donor cell to a recipient cell and expressed in the latter (e.g. Sarand et al. 2000). In contrast to vertical gene transfer, whereby the hereditary lines of most eukaryotic organisms are established, HGT normally occurs between closely-related organisms, but also may occur between bacteria that are distantly related (van Elsas et al. 2006). In a recent study, Agrobacterium tumefaciens was inoculated into clay loam soil with five aminobacter strains and one arthrobacter strain that contained a methylcarbamate-degrading gene. After 61 days of incubation, the A. tumefaciens strain acquired the ability to
degrade carbofuran (Desaint et al. 2003). With regard to physiology, soil bacteria usually possess a wide range of abilities that allow them to rapidly adapt to various chemical conditions and to show physiological responses to new environmental conditions. For example, in soil, *Bacillus subtilis* has two-component regulatory systems that enable it to actively integrate into new environment (Kirby 2009, Mitrophanov and Groisman 2008). Generally, to directly cope with the complex conditions presented by the soil environment, many bacteria fine-tune their genome over evolutionary time. Recent studies on whole bacterial genomes have clearly indicated that bacterial genomes can be relatively large, and the largest bacterial genomes known so far belong to soil bacteria. For example, *Sorangium cellulosum* possesses a genome of 13.1 Mbp in size (Schneiker et al. 2007). Such genomes invariably contain a considerable number of open reading frames with as yet unknown functions. It seems that these unknown putative functions can be employed when bacteria are struggling for survival under fluctuating conditions in soil (van Elsas et al. 2006). Moreover, in soil, a high diversity of potential functions has been revealed by the metagenomic analysis of microbial communities (Rondon et al. 2000, Tringe et al. 2005).

Traditional cultivation-based studies have taught us that a wide range of bacteria belonging to several bacterial phyla inhabit soil. The frequently-isolated groups are Proteobacteria, Firmicutes and Actinobacteria. In addition, bacterial groups identified by molecular approaches, but with fewer cultured representatives, include Planctomycetales, Verrucomicrobiales and Acidobacteria (Buckley et al. 2003, Hugenholtz et al. 1998). Recently, researchers have made improvements to cultivation methods and strategies. For example, the use of polymers as carbon sources and application of longer incubation times and lower incubation temperatures has resulted in the isolation of many novel soil bacteria (Sait et al. 2002). Another example is the modification of standard methods that closely resemble soil environmental conditions. The new approach of high-throughput cultivation based on dilution to extinction can also generate novel isolates from soil (Zengler et al. 2002). However, only a minor fraction of the bacterial diversity has been unlocked so far.

At present, by analyzing 16S rRNA gene sequence similarity, 52 phyla in the domain of Bacteria have been recognized based on data obtained from isolates and on those obtained with direct molecular evidence (van Elsas et al. 2006). However, approximately half of the 52 phyla have not been isolated or cultured. Direct molecular analysis of different soils from boreal environments, using 16S rRNA gene clone libraries, indicates that most soil bacteria fall into several major phylogenetic groups, such as the Proteobacteria, Acidobacteria, Actinobacteria, Verrucomicrobia, Firmicutes and Bacteroidetes (Lesaulnier et al. 2008, Liebner et al. 2008, II).

1.2.2 Fungi

Fungi are primarily aerobic heterotrophs that play an essential role in many soil processes, such as the cycling of minerals and decomposition of organic matter, thus
influencing soil fertility, plant health and nutrition. They also influence plant diversity and functioning and the structure of soil ecosystems. Along with plant and animal kingdoms, fungi represent one of three major evolutionary branches of multicellular organisms, and their uniqueness is reflected by the kingdom of Mycota. Fungal diversity is estimated to be tremendously high. Although only about 75,000 species have been documented so far, it is likely that this represents only 5% of the total number of fungal species, which is predicted to be approximately 1.5 million (Hawksworth 2001). For many years, fungi have been divided into 4 phyla: Ascomycota, Basidiomycota, Chytridiomycota and Zygomycota. Along with the development of molecular phylogenetic analyses, in 2001 a monophyletic clade of Zygomycota was defined as the fifth phylum of fungi, Glomeromycota (Schüssler et al. 2001). In soil, most of the fungi are closely associated with plants as saprotrophs, symbionts or parasites (Finlay 2005).

Saprophytic fungi

Saprophytic fungi have a key role in organic matter decomposition and nutrient recycling. They break down and recycle plant cell walls, which are mainly composed of cellulose and hemicelluloses. Fungi also play a unique role in the degradation of lignocellulose, which is cellulose compounded with lignin (Finlay 2006). The decomposition process usually includes a series of well-characterized stages. In the initial stage, weak parasites and pathogens resident in the living plants are replaced by pioneer saprotrophic fungi that can utilize soluble substrates and storage compounds, but not the structural polymers. After that, pioneer saprotrophic fungi are replaced by polymer-degrading fungi, mainly ascomycetes and basidiomycetes (van Elsas et al. 2006). In the final stage, cellulose is degraded and basidiomycetes take the dominant position, as this group of fungi is able to break down lignin polymers (Steffen et al. 2000). Fluctuations in temperature and moisture greatly affect the decomposition process, and the quantity and quality of organic matter modifies the stages (Murphy et al. 1998, Wilkinson et al. 2002, Aneja et al. 2006, 2007). In addition to true saprotrophs, facultative mycorrhiza-forming fungi are also able to degrade litter or parts of it to a certain extent.

Ectomycorrhizal fungi

Mycorrhizal symbiosis is a common association between plant roots and fungi. The fungus captures nutrients effectively from the soil and translocates part of them to the host plant. In return, the host plant supports mycorrhizal fungi by delivering photosynthesized carbohydrates. The majority of terrestrial plant roots are colonized by symbiotic fungi forming mycorrhizas (van der Heijden and Sanders 2003). Mycorrhizal associations can be divided into seven basic types depending on their morphological features and the fungal and plant species involved (Smith and Read 2008).

EcM is a common mycorrhiza type that is characterized by intercellular fungal colonization and structures such as the mantle and Hartig net. Ectomycorrhiza forming fungi mainly belong to the Basidiomycetes and some to the
Ascomycetes and Zygomygetes (Smith and Read 2008). Ectomycorrhizas are typical for the short roots of woody plants, such as members of the Pinaceae, Betulaceae, Fagaceae and Salicaceae, and are prevalent in northern boreal and temperate forests. Boreal forest soils are generally poor in easily accessible nutrients (Tamm 1991), which may explain why nearly all short roots of boreal forest trees are colonized by ectomycorrhizal fungi (Taylor et al. 2000).

In boreal forest soil, the extraradical mycelium of EcM fungi may form as much as 80% of the total EcM fungal biomass (Wallander et al. 2001). An extraradical mycelium is comprised of hyphae and rhizomorphs, and enlarges the area for nutrient absorption by colonizing areas around roots and by reaching soil pores that are too small for the roots to enter. Rousseau et al. (1994) estimated that a tree root with extending EcM mycelia can exploit a 1000-fold larger soil volume than a non-mycorrhizal root. In terms of the amount and growth of hyphae and rhizomorphs, EcM fungi were classified into different exploration types, which may represent distinct foraging strategies: a contact exploration type in the genera Lactarius, Russula and Tomentella, a short-distance type in Cenococcum and Tylospora, a medium-distance type in Amphinema, Dermocybe, Piloderma and Thelephora and a long-distance type in Boletus and Paxillus (Agerer 2001).

Diversity of EcM

The diversity of ectomycorrhizal communities is high. It is estimated that around 5000–6000 fungal species may be involved in EcM symbiosis, which has evolved independently among multiple lineages of fungi (Hibbett et al. 2000). Phylogenetic studies have revealed that, for example, Agaricales, which include the Boletaceae and Russulaceae, are derived from wood-rotting fungi (Moncalvo et al. 2000), explaining why EcM fungi can express a range of saprotrophic capabilities. Many EcM fungi are able to degrade organic compounds and many EcM fungi inhabit decaying organic matter (Lindahl et al. 2002, Read and Perez-Moreno 2003). As reviewed by Steffen (2003), mycorrhizal and saprotrophic fungi are not fully distinguishable, since their functions and habitats partly overlap. Moreover, many EcM fungi, such as Tomentelloid fungi, were earlier categorized as saprotrophs (Kõljalg et al. 2000). Therefore, the estimation of EcM richness can be extended to a more accurate figure that may be as many as 10 000 species (Taylor and Alexander 2005).

Those EcM fungal species that form easily recognizable epigeous fruiting bodies are relatively widely explored and well known, at least in the boreal and temperate zones. The utilization of DNA methods for studies on EcM diversity has been emphasized since the 1990s, and a large proportion of the EcM fungi that produce inconspicuous hypogeous or resupinate fruiting bodies has been revealed (Horton and Bruns 2001). These fungal species are affiliated, for example, with species in the genera Tylospora, Piloderma, Amphinema and Tomentella in boreal forests (Tedersoo et al. 2003, Taylor et al. 2000, Kõljalg et al. 2000). Although DNA tools are nowadays essential for EcM fungal taxonomic analysis, fungal DNA
sequences from environmental samples often remain unidentified due to the lack of a match in databases, showing that only documented EcM fungal phylotypes can be recognized.

Numerous factors affect the diversity of an EcM community in forest soil, such as the species, age and vigour of the host trees, edaphic and environmental conditions, the availability of fungal inoculum, competition, the microflora and microfauna, anthropogenic stress, site history, habitat size and the degree of isolation (Deacon and Fleming 1992, Erland et al. 1999, Peay et al. 2007). There is also temporal and spatial variation in the structure of EcM communities. The dynamics of EcM communities make them challenging to sample, as a single sampling event represents the EcM community only at a particular moment. Temporal variation in EcM fungal communities has been studied (Koide et al. 2007) but it is still poorly known. The spatial heterogeneity of EcM communities has also gained considerable attention, and studies have demonstrated that various EcM fungi follow variable distribution patterns (Tedersoo et al. 2003, Koide et al. 2005, Genney et al. 2006).

**Studying EcM**

Despite its ecological relevance, the study of EcM mycelia in situ is challenging, since it has been impossible to distinguish EcM hyphae from saprotrophic hyphae or other fungi growing in the forest soil. In terms of molecular methods, the composition of EcM fungal species can be determined by extracting fungal DNA from soil and by using sequencing or fingerprinting techniques for identification (Azul et al. 2010, Cavender-Bares et al. 2009, Hynes et al. 2010). However, due to the polyphyletic property of EcM fungi, it is not possible to develop specific primers for the investigation EcM fungi.

In 2001, in-growth mesh bags, filled with acid-washed quartz sand, were introduced to separate EcM mycelial growth in the field from other types of fungi, for example saprotrophs (Wallander et al. 2001). The small mesh size of the bags prevents penetration of plant roots, meanwhile the poor nutrient sand provides colonization priority to EcM fungi, since EcM fungi absorb carbon released by host trees, while the saprophyses cannot. It has been documented that after one growing season, EcM fungi are the main colonizers in the in-growth bags and the proportion of non-mycorrhizal fungi is minor (Wallander et al. 2001, Kjøller 2006).

### 1.2.3 Diversity and function

The importance of taxonomic diversity for ecosystem function was recognized more than half a century ago by MacArthur (1955), who concluded that increasing diversity should increase system stability because of the redundancy of species functions. Fungal diversity may positively correlate with the rate of decomposition under changing environmental conditions (Toljander et al. 2006), and the diversity of decomposers improves resilience and resistance (Griffiths et al. 2000). Species that are functionally similar, but react to environmental changes differently may increase stability, but only in certain circumstances (Chapin et al. 1997). These
finding indicate that there is a relationship between ecological diversity and function. However, Bell et al. (2005) demonstrated that the relationship between taxonomic diversity and ecosystem functioning is not linear, but asymptotic in form. Such a relationship suggests a redundancy theory that ecosystem functioning may not change further with altered species diversity. In other words, the number of species needed to make a system stable and the amount of damage the ecosystem can withstand without losing its functionality are not known and are difficult to determine.

With approximately one million bacteria per gram of soil (Gans et al. 2005), functional redundancy is most likely so high that losing even a large part of the taxonomic diversity should not affect the functional properties of the whole community. In fact, this redundancy may be responsible for the sensitivity of minor groups in the community. For example, according to recent estimates, the decrease in taxonomic diversity in metal-polluted soils can reach 99.9%, with most of the eliminated species belonging to rare taxa (Gans et al. 2005). In addition, based on data from high-resolution fingerprinting techniques, the diversity of prokaryotic strains was found to reflect the habitat from where these originated, and soil management and other features markedly influenced the community composition (Kurenbach et al. 2003, Ashelford et al. 2000).

1.3 Sequencing methods in microbial diversity studies

Earlier, organisms that could not be cultured were not studied at all unless they posed some significant known threat to human health or commerce. Now however, thanks to the recent advent of rapid molecular sequencing and analysis technology, it is possible to collect samples from the environment and directly survey their diversity without having to culture them. During the past two decades, culture-independent analysis with ribosomal RNA genes has been widely utilized to study archaeal, bacterial and eukaryotic organisms from various habitats. Generally, the scope of microbial diversity according to these studies has been far higher than previously estimated based on cultivation analysis, including numerous novel lineages within the described phyla, and many identified novel phyla (Youssef and Elshahed 2008).

1.3.1 rRNA genes

The study of microbial diversity appears to have become increasingly popular using the ribosomal RNA gene sequence as a tool to assess the evolutionary relationships among microbes in the environment. The genes coding for the bacterial small and large ribosomal subunits, 16S and 23S rRNA genes, respectively, and fungal 18S and 28S rRNA genes, contain both conserved and variable regions. The differences in these genes can be used to infer the relationships between DNA sequences from different species (Van de Peer et al. 1996). Universal primers have been developed to amplify the ribosomal RNA gene regions from bacteria and fungi (Edwards et al. 1989, White et al. 1990). Fungal internal transcribed spacer region (ITS), located between the 18S and 28S rRNA gene subunits, is usually used for fungal species level identification, because
the 18S subunit sequence is not as diverse as the ITS (Iwen et al. 2002).

Due to the superior sensitivity of PCR, it has become unnecessary to cultivate microbes as the rRNA gene can be directly amplified from samples. Sequences obtained from the environment by direct amplification were described as the sole means to provide information for at least 99% of the prokaryotes in natural communities (Amann et al. 1995). This approach has been so popular, that the public data base INSDC (i.e. DDBJ/EMBL/GenBank, URL: http://www.insdc.org) mostly includes rRNA gene sequences amplified directly from environmental samples without prior culturing.

1.3.2 Sanger DNA sequencing

Sanger DNA sequencing technology (Sanger et al. 1977) has been applied in the identification of microbes since the 1990s (Fredricks and Relman 1996). Compared with traditional morphological and chemical (for example phospholipid-derived fatty acids, PLFA) identification, this technology is more precise. In microbial ecology studies, Sanger DNA sequencing technology has been used, for example, in identifying cultivable bacteria (Heylen et al. 2006), fungal sporocarps (Niskanen et al. 2008) and ectomycorrhizal root tips (Twieg et al. 2007). Nevertheless, when studying the total diversity of a community, the above technology is quite time consuming. It is not suitable for a large number of samples or total diversity analysis. DGGE (denaturing gradient gel electrophoresis) and SSCP (singles-strand conformation polymorphism) are often used in assessing microbial diversity. It is also possible to sequence the bands for phylogenetic information, but these technologies only identify the key species with a high abundance (Brettar et al. 2006), and numerically rare phylotypes are generally not detected (Bent and Forney 2008). Moreover, a single DGGE band may include several types of ribosomal RNA gene sequences (Costa et al. 2006). To separate different sequences, DNA cloning technology can be utilized.

1.3.3 Cloning and sequencing

DNA cloning is the process of moving a gene from the chromosome it occurs to an autonomously replicating vector. In microbial ecology studies, cloning and sequencing is a rapid approach to reveal microbial diversity in the environment, especially with the assistance of modern robot technologies. Studies conducted using this technology usually provide a few thousands of sequences (Partanen et al. 2010, Hultman et al. 2010). The number of sequences in clone libraries has been steadily increasing, especially with the decreasing sequencing costs, and the recent availability of sequence analysis programs and fast growing databases (Youssef et al. 2008). However, cloning-based approaches have certain disadvantages. First, it is challenging to obtain a representative sample to study the diversity and abundance of microbes in the sample (Ranjard et al. 2003). It is also difficult to estimate the number of clones that need to be sequenced to reach the desired coverage in a sample. Microbial groups that are abundantly present in samples overshadow the ones present in lower amounts that might not be sampled and sequenced (Curtis and Sloan 2005).
magnitude of microbial diversity is far too large to be understood and it is demanding to sequence every microbe in a sample. DNA cloning and sequencing methods currently require large amounts of funding, although prices are decreasing.

### 1.3.4 Pyrosequencing

In recent years, several next-generation sequencing techniques have been developed, such as pyrosequencing first used for genome sequencing (Margulies et al. 2005). Massive pyrosequencing can presently produce more than one million high-quality reads per run and read lengths ranging from 100 to 400 bases (http://www.454.com). This allows sequencing of the entire genome (van Schaik et al. 2010) or selectively variable short domains of the rRNA gene (McLellan et al. 2010), apparently with good accuracy (Huse et al. 2007). Recently, Stach and Bull (2005) concluded that clone libraries of hundreds of samples are extremely unlikely to produce accurate estimates, until very large libraries can be investigated, indicating how quickly the situation has changed.

During the past three years, a vast quantity of sequence data has been generated by pyrosequencing. The most recent publications on microbial communities, employing the pyrosequencing method, are presented in Table 1. The numbers of sequences reported in the papers reflect a revolution in molecular methods for studying microbial ecology. The comparison of microbial communities, especially communities with a high number of sequences and in many different environments, is a challenge to researchers.

The traditional sequence assignment programs are either too slow (for example, previous versions of the Staden Package) or lack the functions to simultaneously analyze separate samples (for example SONS and DOTUR). Authors use different means to analyze pyrosequencing data, but the available means are mostly only suitable for certain purposes. Therefore, the development of a program that can analyze microbial community data and fulfil case-specific demands is essential.

### 1.3.5 Diversity estimates

When accessing less abundant members of the community, the increase in sequence datasets produced by DNA cloning or pyrosequencing allows the utilization of various statistical approaches in evaluating microbial diversity (Youssef et al. 2008, Jumpponen and Jones 2009). Regardless of the methods and materials utilized for microbial community study, it has been observed that the predicted community richness value is dependent on the size of the clone library used in the calculation (Dunbar et al. 2002). Adequate sampling to encounter all bacterial species in the sample could theoretically be a solution for this problem, but the majority of bacterial species are present at a low frequency in the complex microbial community making an exhausting analysis virtually impossible. The detection of all bacterial species in a complex ecosystem (for example soil) remains a questionable task even with the recent advanced 454 pyrosequencing technology.
Table 1. Review of recent bacterial and fungal community studies using 454 pyrosequencing

<table>
<thead>
<tr>
<th>Sampling resource</th>
<th>Organism</th>
<th>Target gene</th>
<th>Total No. of sequences</th>
<th>Length of sequences$^{1)}$ (bp)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil</td>
<td>Bacteria</td>
<td>16S</td>
<td>3467</td>
<td>101</td>
<td>Barriuso et al. 2010</td>
</tr>
<tr>
<td>Western Arctic ocean</td>
<td>Bacteria</td>
<td>16S</td>
<td>250555</td>
<td>62</td>
<td>Kirchman et al. 2010</td>
</tr>
<tr>
<td>Arctic ocean</td>
<td>Bacteria</td>
<td>16S(V6)</td>
<td>313827</td>
<td>79</td>
<td>Galand et al. 2010</td>
</tr>
<tr>
<td>Glacier</td>
<td>Bacteria</td>
<td>16S</td>
<td>234000</td>
<td>100</td>
<td>Schütte et al. 2010</td>
</tr>
<tr>
<td>Baltic Sea</td>
<td>Bacteria</td>
<td>16S(V6)</td>
<td>161600</td>
<td>79</td>
<td>Andersson et al. 2010</td>
</tr>
<tr>
<td>Sewage from wastewater treatment plant</td>
<td>Bacteria</td>
<td>16S(V6)</td>
<td>215090</td>
<td>43</td>
<td>McLellan et al. 2010</td>
</tr>
<tr>
<td>Western English Channel</td>
<td>Bacteria</td>
<td>16S(V6)</td>
<td>182560</td>
<td>79</td>
<td>Gilbert et al. 2009</td>
</tr>
<tr>
<td>Marine</td>
<td>Bacteria</td>
<td>16S(V3)</td>
<td>9643</td>
<td>155</td>
<td>Alvarez et al. 2009</td>
</tr>
<tr>
<td>Biofilm</td>
<td>Bacteria</td>
<td>16S(V6)</td>
<td>119480</td>
<td>79</td>
<td>Parameswaran et al. 2010</td>
</tr>
<tr>
<td>Deep-sea clams</td>
<td>Bacteria</td>
<td>16S(V5)</td>
<td>24606</td>
<td>170</td>
<td>Stewart and Cavanaugh 2009</td>
</tr>
<tr>
<td>Soil</td>
<td>Bacteria</td>
<td>16S(V6)</td>
<td>17471</td>
<td>101</td>
<td>Barriuso et al. 2010</td>
</tr>
<tr>
<td>Ocean</td>
<td>Bacteria</td>
<td>16S(V6)</td>
<td>250555</td>
<td>62</td>
<td>Kirchman et al. 2010</td>
</tr>
<tr>
<td>Soil</td>
<td>Bacteria</td>
<td>16S(V1-V2)</td>
<td>31373</td>
<td>150-316</td>
<td>Koopman et al. 2010</td>
</tr>
<tr>
<td>Arctic Ocean</td>
<td>Bacteria</td>
<td>16S(V6)</td>
<td>740353</td>
<td>&gt;50</td>
<td>Galand et al. 2009</td>
</tr>
<tr>
<td>Soil</td>
<td>Fungi</td>
<td>ITS(ITS1)</td>
<td>29910</td>
<td>&gt;200</td>
<td>Jumpponen et al. 2010</td>
</tr>
<tr>
<td>Salt-marsh plants</td>
<td>Fungi</td>
<td>ITS</td>
<td>5199</td>
<td></td>
<td>Gillevet al. 2010</td>
</tr>
<tr>
<td>Soil</td>
<td>Fungi</td>
<td>ITS(ITS1)</td>
<td>166350</td>
<td>&gt;100</td>
<td>Buée et al. 2009</td>
</tr>
<tr>
<td>Soil</td>
<td>Fungi</td>
<td>ITS(ITS1)</td>
<td>84956</td>
<td>268</td>
<td>Jumpponen and Jones. 2010</td>
</tr>
<tr>
<td>Plant</td>
<td>Fungi</td>
<td>SSU$^{2)}$</td>
<td>179279</td>
<td>160-279</td>
<td>Öpik et al. 2009</td>
</tr>
<tr>
<td>Soil</td>
<td>Fungi</td>
<td>ITS(ITS1)</td>
<td>18020</td>
<td>265</td>
<td>Jumpponen and Jones 2009</td>
</tr>
</tbody>
</table>

1) approximate value or range

2) ssu=small subunit ribosomal RNA gene of Glomeromycota
1.3.6 Other DNA methods

As a consequence, many researchers have developed or used alternative methods. DNA arrays that consist of a series of DNA probes (reporters) immobilized on a solid substrate such as a glass slide or a silicon chip have an enhanced capacity for achieving high levels of quality control and are easy to use. DNA arrays have two main uses in microbial ecology: detection of known microbes and the assessment of diversity (Hultman et al. 2008). In general, a PCR amplification step using specific primer sets is used to generate labeled DNA (Berthet et al. 2008). This requirement of predesigned oligomers for amplification and hybridization opens up the risk that many unknown organisms will remain undetected.

1.4 Heavy metal contamination of soils

Heavy metals are natural constituents of the Earth's crust, but human activities have largely altered their geochemical distribution and biochemical balance in the soil (Giachetti et al. 2006). It is commonly accepted that heavy metals strongly influence bacterial communities, although bacteria obviously adapt to heavy metal contaminated environments by developing tolerances, as reviewed by Suhadolc et al. (2004). Heavy metals accordingly affect some processes in soil such as the degradation of organic pollutants, enzyme activities, the mineralisation of organic matter and nitrogen transformations (Suhadolc et al. 2004).

1.4.1 Lead and lead pollution

Lead (Pb) is a widespread, soft metal that has been commonly used by humans for a long time and for a variety of purposes, such as in building construction, lead-acid batteries, bullets and shot. To prevent "engine knocking" in high compression internal combustion engines, lead has been used as a gasoline additive for a century. The possession and use of leaded gasoline in a regular on-road vehicle is prohibited in most countries in the world. However, fuel containing lead may continue to be sold for off-road uses, including aircraft, racing cars, farm machines, and boat engines.

In general, Pb mobility is low because of its low solubility. The low solubility may further result in low bioavailability. When a Pb pellet enters the soil, sorption and ion exchange processes start with the soil's solid phase. The transformation of metallic Pb from pellets in soil can be described as: \( P \rightarrow PbO \rightarrow PbCO_3/Pb_3(CO_3)_2(OH)_2 \) (Ma et al. 2007). Elemental Pb is transformed through different phases including oxidation, carbonation and hydration, and become diversified into dissolved and particulate Pb species. The newly formed oxidized Pb phases mainly include cerussite, hydrocerussite, possibly anglesite, and massicot/plattnerite (Lin et al. 1995, Cao et al. 2003). The Pb carbonates formed create a protective layer preventing further weathering of the pellets, which turns the Pb bullet into a two-layer system, with the inner layer being predominantly massicot, and the outer layer predominantly hydrocerussite. Weathering of Pb bullets can increase Pb bioavailability in contaminated soils, but this process is affected by variety of
environmental factors. For example, the weathering rate of Pb is high in organic soils (Dermatas et al. 2006). Both Pb solubility and the transformation of Pb into oxidized species are related to soil pH (Cao et al. 2003).

Lead contamination in the environment is of particular concern, as it is a known toxin. Lead toxicity to humans and wildlife has been extensively studied. The greatest risk emanates from the direct ingestion of bullets, contaminated soil or plants (Vantelon et al. 2005). In humans, lead can cause blood and brain disorders, damage nervous connections and reproductive function, and lead in soil may also have direct toxic effects on plants and aboveground fauna (Singh et al. 1997).

Due to the close contact with the surrounding soil, the soil microbial community can be expected to be the first and most seriously affected by lead contamination of the soil. Pb can affect the growth, morphology and metabolism of soil microorganisms through functional disturbance, protein denaturation, and the destruction of cell membrane integrity (Leit et al. 1995). In soil exposed to Pb, the amount of bacteria and fungi decreased with the incubation time, and the bacterial number diminished sharply (Wang et al. 2010). Possible lead-induced changes in belowground food webs can be expected to affect the aboveground biota, and possibly alter the functioning of the entire ecosystem through the interactions between aboveground and belowground biota (Rantalainen et al. 2006).

1.4.2 Shooting ranges in Finland

A controlled environment for practicing shooting is safer. Shooting ranges provide such an environment for the discharge of firearms. In Finland, more than 500 tons of Pb is produced each year for shotgun ammunition and bullet jackets (Sorvari et al. 2006). An average shot contains up to 97% metallic lead, 2% antimony, 0.5% arsenic and 0.5% nickel, while jacketed bullets contain up to 90% metallic lead, 9% copper and 1% zinc (Sorvari et al. 2006). There are abundant amounts of lead deposited in soils at shooting ranges, as in Finland the annual discharge of lead can be from 120 to 15 000 kg (Sorvari et al. 2006), and shotgun pellet accumulation may be as high as tens of thousands of kilograms per hectare per year (Darling and Thomas 2003). Due to the large number of discharged Pb bullets, shooting ranges have been identified as one of the main sources of soil pollution with Pb. Sorvari and coworkers reported that shooting ranges account for 5% of the total number of contaminated areas in Finland (Sorvari et al. 2006). Most Finnish outdoor shooting ranges are situated in sandy esker areas, which are very often important aquifer areas, commonly used as domestic water sources (Sorvari et al. 2006). Corrosion of the released Pb bullets is continuous and ongoing, and in such areas could release significant quantities of Pb into the drinking water supply. Those shooting ranges located deep in forest areas cannot be considered less environmentally dangerous, as Pb is toxic to the forest ecosystem as well as to human beings. Lead may also be transported to ground water from the soil (Murray et al. 1997), thus threatening the health of human populations.
Compared to lead contamination originating from industrial sources, that caused by shooting range activities has received much less attention (Lin et al. 1995, Darling and Thomas 2003). Shooting range lead may constitute a serious threat to various organisms and ecological risks may also be significant. Several previous studies have reported high levels of lead and associated effects on both local and migratory aquatic and terrestrial biota (Sorvari et al. 2006). In a shooting range, although the area of contamination is often limited, the coverage is always large, making it necessary for a soil survey to be comprehensive. A heavily contaminated area is always formed behind the stop butts, with maximum soil Pb concentration ranging up to 150 000 mg kg\(^{-1}\) (Scheetz and Rimstidt 2008), while the average natural background concentration of Pb in soils is typically in the order of 10–30 mg kg\(^{-1}\) (Adriano 1986). The long residence time of lead also makes it a potential long-term environmental hazard or risk. Although most of this lead may remain in the pellets, even for centuries (Jørgensen and Willems 1987), part of it is continuously transformed into soluble forms that have bioaccumulation potential (Hui 2002, Labare et al. 2004).

2 AIMS OF THE STUDY

Shooting ranges represent important point sources of lead contamination, which makes them ideal model systems to study the effects of lead contamination on soil microbial communities. Although there has been wide interest in microbial diversity under contamination by heavy metals, the effects of lead on the composition of microbial communities has remained poorly known, especially in the boreal environment.

The general aim of this study was to characterize the microbial communities in lead-contaminated an uncontaminated boreal forest soil. The specific aims of this study were to:

1. Describe the bacterial and fungal communities in lead contaminated and uncontaminated soil (I, II, III);
2. Estimate the bioavailability of lead and effects of lead on soil biota, microbial diversity and species composition (I);
3. Gain a detailed view of the changes in microbial community structure (II) and in the communities of EcM fungi (III) due to lead contamination; and
4. Develop a computer program to manipulate large volumes of DNA sequence data and calculate general OTU series with frequency values from a large number of samples.
3 MATERIALS AND METHODS

3.1 Hälväälä shooting range

Healthy looking, well growing trees are interpreted as a sign of a clean, unpolluted environment. A good example is the forest at a shooting range, heavily contaminated by lead from shotgun pellets but showing no visible sign of stress. At the Hälväälä shooting range, located in Southern Finland (61°00'N, 25°80'E), the model ecosystem appears normal from a distance, although the ground in the most heavily contaminated sections is virtually covered with lead pellets from many years of shotgun activity. Even when the pellets are removed, the total concentration of lead in soil is high and the ecosystem can be considered as saturated. The conditions at Hälväälä are apparently such that at least pine growth and cover vegetation are not visibly disturbed (Rantalainen et al. 2006).

One sector of the Hälväälä shooting range, hereafter referred to as the old contaminated area (OC), was used as a shot gun shooting range during the years 1964-1987, after which was allowed to reforest naturally. The shooting sector was then transferred to a new direction, and this sector is still in use (new contaminated area, NC). A medium contaminated area (M), which is close to the OC, was selected and earlier described together with the OC (Turpeinen et al. 2000, Salminen et al. 2002, Rantalainen et al. 2006). A comparable nearby uncontaminated area was chosen as a control (clean control, C). A detailed map of Hälväälä shooting range is presented in III.

3.2 Sampling and methods

As illustrated in Figure 1, the forest humus layer soil, pine root tips, EcM sporocarps and in-growth bags were sampled in 2003-2007 (Table 2). The sampling procedures and experimental methods used in this study are mainly described in detail in the original publication, and the methods are summarized in Table 3.

Table 2. Sampling of this study

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>II</td>
<td>I, III</td>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>M</td>
<td>II</td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>OC</td>
<td>II</td>
<td>I, III</td>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>NC</td>
<td>II</td>
<td>I, III</td>
<td></td>
<td>III</td>
</tr>
</tbody>
</table>
Figure 1. DNA sequencing methods used in the studies of this thesis
Table 3. Methods used in this study

<table>
<thead>
<tr>
<th>Method</th>
<th>Described in study</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA extraction from soil/root tips/in-growth bags</td>
<td>I, II and III</td>
<td>FastDNA SPIN Kit for Soil</td>
</tr>
<tr>
<td>Clone library construction for fungal ITS-region</td>
<td>II</td>
<td>Hultman et al. 2010</td>
</tr>
<tr>
<td>Clone library construction for bacterial 16S rRNA gene</td>
<td>I and II</td>
<td>Partanen et al. 2010</td>
</tr>
<tr>
<td>Sanger DNA sequencing</td>
<td>I, II and III</td>
<td>Sanger et al. 1977</td>
</tr>
<tr>
<td>454 pyrosequencing</td>
<td>III</td>
<td>Jumpponen and Jones 2009</td>
</tr>
<tr>
<td>Measurements of total and water soluble Pb concentration</td>
<td>I</td>
<td>Hartikainen and Kerko 2009</td>
</tr>
<tr>
<td>Isolation and cultivation of soil bacteria</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Shannon diversity index</td>
<td>II and III</td>
<td>Shannon 1948</td>
</tr>
<tr>
<td>Chao 1 diversity index</td>
<td>II and III</td>
<td>Chao 1984</td>
</tr>
<tr>
<td>Richness</td>
<td>II and III</td>
<td>Krebs 1998</td>
</tr>
<tr>
<td>Evenness indices</td>
<td>II</td>
<td>Fisher et al. 1943</td>
</tr>
<tr>
<td>Simpson diversity index</td>
<td>II and III</td>
<td>Simpson 1949</td>
</tr>
<tr>
<td>Nonmetric multidimensional scaling (NMS)</td>
<td>III</td>
<td>McCune and Grace 2002</td>
</tr>
</tbody>
</table>
3.3 A new program for OTU assignment

Introduction of MuSSA (Multi-Sample Sequence Analyzer)

Traditional cloning and sequencing methods, which generate hundreds of sequences have been supplemented by 454 pyrosequencing techniques, which generate thousands of sequences. Sequence data handling with commonly used programs, for example older versions of the Staden Package (Staden 1996) is becoming difficult and exceedingly time consuming and novel tools are needed to handle the vast amounts of data being generated. The field of microbial ecology has recently benefited from the availability of a number of tools for describing and comparing microbial communities using robust statistical methods, including CAP3 (Huang and Madan 1999) and DOTUR (Schloss and Handelsman 2005). However, although these programs provide methods for rapid calculation, it has become necessary to make them more amenable to larger datasets and the numbers of samples

![Figure 2. User's panel of the Multi-Sample Sequence Analyzer (MuSSA).](image-url)
from pyrosequencing. As part of this thesis a new program was developed, MuSSA (Multi-Sample Sequence Analyzer), to handle sequence data from novel high throughput sequencing approaches in microbial community analyses. MuSSA uses 454 sequencing Fasta files as an input. From this input unique identifiers are generated based on the primer-incorporated DNA tags that are used as the sample level identifier. The program pipeline then calculates operational taxonomic units (OTUs) based on user-determined similarity thresholds. The frequency of these OTUs and their distribution across the samples can then be parsed back to the sample level, for which it is possible to calculate community summary statistics such as diversity, richness and evenness indices. The greatest advantage of the program is that large volumes of sequence data can be manipulated and general OTU series with frequency values can be calculated from a large number of samples.

**Working processes and Functions**

The user interface of the program is as illustrated in Figure 2, demonstrating that MuSSA is capable of processing the following functions in the order of working processes:

A. Sequence quality control (SQC) and input file
With the SQC function, one can set cut-off values for removing sequences that are short, with uncertain base pairs and missing primers or tags. SQC can also be used independently for normal sequence dataset editing. The input file must be in Fasta format (EMBL).

B. Primer and tag identification
Fasta formatted raw data include all DNA sequences that are generated by 454 pyrosequencing, and these can be directly handled by MuSSA, since MuSSA is able to recognize the primers and identification tags (designed for sample identification) of DNA sequences. Sequences are categorized according to the samples and those sequences without a tag or primer are omitted.

C. Assigning sequences to general OTU series among multiple samples
If thousands of DNA sequences are obtained within one sample when examining the microbial diversity, efficient programs, for example, DOTUR can be employed for OTU assignment. When the sequences are obtained from two samples, the same method can be applied, but note that there is no information that correlates the two OTU series between the two samples. The information on the phylogenetic positions of the OTUs in these two samples is not automatically comparable. With larger numbers of samples, the situation is even more complicated. As the main and novel function of the MuSSA program is to calculate the general OTU series among all samples, instead of independent OTUs in each sample. It allows the user to obtain a comprehensive view of the phylogenetic distribution in the whole dataset. The similarity value of OTU assignment can be freely set.

D. Calculation of OTU frequency in samples
In the output files of MuSSA, there is a table matrix that counts the number of sequences of each OTU in different samples (see an example Table 4). The numbers describe the
frequency of OTUs in each sample and they form the basis for rarefaction analysis and the calculation of diversity indices (for example Chao1, Shannon and Simpson indices).

E. Output files
The output files include a table of the frequency matrix and a list of representative sequences from all OTUs. MuSSA is still under development and the next version will contain more output files, for example rarefaction figures and files that are directly utilizable as input files by EstimateS (Colwell 2006).

Properties and system requirement

MuSSA is an extension of CAP3 (Huang and Madan 1999). MuSSA can analysis numbers of samples and generate a OTU series, while Cap3 cannot. Additionally MuSSA has also the sequence quality control system. In 2009, MuSSA was powered by Java SE Development Kit, Version 6 (http://www.oracle.com). The newest version is 1.5.

Table 4. A output table of MuSSA.

<table>
<thead>
<tr>
<th>OTU001</th>
<th>Sample1</th>
<th>Sample2</th>
<th>Sample3</th>
<th>Sample4</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTU002</td>
<td>12</td>
<td>5</td>
<td>3</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>OTU003</td>
<td>3</td>
<td>0</td>
<td>254</td>
<td>254</td>
<td></td>
</tr>
<tr>
<td>OTU004</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>OTU005</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>OTU006</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>OTU007</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>OTU008</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>…</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Software requirement:

Windows XP (32/64-bit); a Pentium IV 2.0 GHz or faster processor; 1G or more of RAM (depending on the size of the sequence dataset, more RAM maybe needed. 6G or more is recommended); JAVA Runtime Environment 1.50 or later release (freely downloadable via http://www.oracle.com/); and 100 GB hard drive space.

The program has not been published or released for public use, but has only been applied in our research group. It was designed to fulfill our own research targets, and in its present form is probably not suitable for general use. Both sequence data generation on the 454 platform and its manipulation to generate numerical data from raw Fasta files are rapidly evolving and widely spreading applications in microbial ecology, meaning that other groups develop similar programs for different purposes such as MOTHUR (Schloss et al. 2009).
4 RESULTS AND DISCUSSION

4.1 Microbial communities in Pb contaminated soil

Heavy metals have been reported to reduce forest productivity (Pukacki and Kaminska-Rozek 2002), inhibit growth of trees, especially seedlings (Menon et al. 2007, Fedorkov 2007, Bojarczuk and Kieliszewska-Rokicka 2010), and negatively affect biodiversity and stand structure (Koptskik et al. 2004, Chernenkova and Kuperman 1999). Although the forest soil in the Hälvälä area is heavily contaminated by lead, the visual appearance of the forest at the shooting range is indistinguishable from the nearby clean forest and at least pine growth and ground cover vegetation are not visibly disturbed (Rantalainen et al. 2006). As summarized in Table 5, the only measurable difference between the clean and contaminated areas is the lead concentration in the soil. Heavy metals in soil exhibit toxic effects on soil microbes, which may reduce microbial quantities (both biomass and diversity) and their level of activity (Khan et al. 2010, Wang et al. 2010). Bacterial Pb tolerance is generally thought to be lower than in fungi (Rajapaksha et al. 2004). Both fungal and bacterial biomasses were negatively affected by lead in earlier investigations in our study area at the Hälvälä shooting range (Rantalainen et al. 2006), but the bacterial biomass was shown to decrease less than the fungal biomass in the lead contaminated areas at Hälvälä (Salla Selonen, unpublished data). Both observations suggest that the bacterial community was less influenced by lead than the fungal community. Total fungal richness (number of observed OTUs, $S_{obs}$), diversity (Simpson’s D and Shannon H) and extrapolated species richness (Chao1) were also altered when comparing the clean and lead contaminated areas (II), indicating that the total fungal community was affected by Pb contamination. The richness and diversity indices of the bacterial community measured from the same template DNA remained unchanged under Pb contamination (II). This contradicts the findings of Khan et al. (2010), who reported that in a short-term exposure to lead, bacteria were more sensitive than fungi when studying the soil of a maize crop field. Although fungal diversity in agriculture fields is totally different from that in coniferous forest soil, in general it seems possible that the fungi responded more slowly than the bacteria, and no change was thus seen in this short-term experiment. This suggestion is supported by the finding that the amount of fungi decreased as a function of incubation time after Pb exposure in soil (Wang et al. 2010). Nevertheless, based on our results, the exposure of a previously uncontaminated site to lead is likely to have a dramatic initial effect on the microbiota, after which the fungal strains will either begin to adapt and acquire resistance traits, or the initially sensitive fungal community will be replaced by lead-resistant microbes.

4.2 Ectomycorrhizal fungi

Although total fungal richness, diversity and extrapolated species richness differed between the clean and lead polluted areas (II), EcM fungal richness (number of OTUs, S), diversity (Fisher’s, Simpson’s D, and Shannon’s H) and extrapolative richness estimators (Chao1, Jackknife) did not vary between clean and contaminated areas (III),
Table 5. Environmental parameters at Hälvälä

<table>
<thead>
<tr>
<th>Area</th>
<th>1) Total Pb concentration mg/kg DW</th>
<th>2) Total Pb concentration mg/kg DW</th>
<th>2) Water soluble Pb concentration mg/kg DW</th>
<th>1) pH (CaCl₂)</th>
<th>1) Biomass Total PLFA μg / g OM of soil</th>
<th>1) Biomass Bacteria PLFA μg / g OM of soil</th>
<th>1) Biomass Fungi PLFA μg / g OM of soil</th>
<th>1) Tree age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>75.6±49</td>
<td>106±88</td>
<td>3.2±1.6</td>
<td>3.25±0.1</td>
<td>1815±437</td>
<td>812±219</td>
<td>97.4±34.4</td>
<td>21.6±2.3</td>
</tr>
<tr>
<td>M</td>
<td>8700±6485</td>
<td>-</td>
<td>-</td>
<td>3.45±0.2</td>
<td>1347±313</td>
<td>632±159</td>
<td>43.7±12.5</td>
<td>19.6±2.3</td>
</tr>
<tr>
<td>OC</td>
<td>18800±15770</td>
<td>42350±8634</td>
<td>110±35.5</td>
<td>3.54±0.3</td>
<td>1116±525</td>
<td>527±238</td>
<td>51.2±24.7</td>
<td>22.7±1.9</td>
</tr>
<tr>
<td>NC</td>
<td>-</td>
<td>26730±8935</td>
<td>120±19.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

1) Rantalainen et al. 2006. All analyses from 2004.
2) I. Total and water soluble Pb analyses from 2005.
- Not measured.
C=Clean area, M=Medium contaminated area, OC=Old contaminated area, NC=New contaminated area.
value = mean±1SD
indicating no differences in EcM communities among the areas with different Pb exposure histories. The different coverage of the fungal community (total fungal diversity in II, EcM fungal diversity in III) may explain this discrepancy, as it is estimated that there are approximately 700 EcM fungal species, 700 wood decomposing fungal species and 500 species living as saprotrophs in soil and litter (Dahlberg 2002). EcM fungi thus represent only about one-third of the species in coniferous forest soils. Other types than EcM fungi may therefore be responsible for the change in total fungal diversity due to contamination, although at the genus level some EcM fungi in II showed a response to lead.

Fungal in-growth bags (Wallander et al. 2001), which were used to investigate EcM diversity in III, primarily target the actively growing non-saprophytic species during the incubation period (Wallander et al. 2001), while fungi in forest humus are comprised of all fungal types, including EcMs (II). In soil, EcM fungi are not as dominant and competitive as they are in in-growth bags. In soil, they suffer not only Pb contamination, but also competition with other, saprophytic species. The in-growth bags, at least at the beginning of the incubation, are lead-free environments without any competitors. Humus samples were taken in September 2005 (II), while in-growth bags were incubated for two growing seasons (2006-2007, III). Thus, the first sampling strategy assessed the community at a certain time point, while the latter strategy collected species for a longer period. In a Scots pine forest, Pickles et al. (2010) analyzed the soil EcM communities by means of morphotyping combined with DNA sequencing and found a marked change in the distribution of ectomycorrhizal taxa between two different sampling times in the same location, which might result from the different living stages of individual mycorrhizal species (Downes et al. 1992) and the renewal of fine roots in forest soil (Högberg et al. 2002, Guo et al. 2008). These findings indicate that temporally different sampling of EcM fungi may yield a different appearance of the communities.

EcM fungi are a critical functional component of many of the world's forest ecosystems (Smith and Read 2008). The symbiotic association between the hosts and EcM fungi plays an important role in both nutrient dynamics (Read and Perez-Moreno 2003) and carbon flow (Högberg et al. 2002). At Hälvälä, no differences were found in tree growth or in the ages or numbers of trees between the clean and contaminated areas (Rantalainen et al. 2006, I, III), indicating that EcM communities are functioning well under lead contamination. Although our 454 pyrosequencing results suggested a shift in EcM communities in the contaminated areas, only three of the ten most frequent genera statistically differed in frequency between clean and contaminated areas. Two of them had decreased and one increased in frequency in the contaminated areas (III), and most of the dominant taxa had not responded to lead in the soil.

Some microcosm and field studies have shown competitive replacement in EcM fungal communities. For example *Pisolithus tinctorius* mycorrhiza was replaced by an unknown mycorrhizal fungus in a "Rhizobox" with one-month-old *Pinus densiflora* seedlings (Wu et al. 1999).
decrease in *Paxillus involutus* was followed by an increasing frequency of *Suillus bovinus* in a laboratory experiment with Scots pine seedlings (Landeweert et al. 2003), and *Rhizopogon occidentalis* colonising *Pinus muricata* seedlings was strongly increased by contamination (Kennedy and Bruns 2005). It has been documented that competitive interaction between fungal species is strongly related to the local conditions (Kennedy et al. 2007). In the Hälvälä case, lead may distinctly impact on the compatibility of some fungi in contaminated areas. Some were enriched and some were reduced in frequency and this might balance the changes in diversity and richness of EcM fungi (III).

### 4.3 Community composition of EcM fungi

Comparing the EcM fungal communities in II and III, a similar fungal genus distribution was observed, as the most common EcM genera in terms of frequency were *Thelephora, Tylospora, Pseudotomentella, Piloderma, Lactarius, Cortinarius* and *Russula*. Although no change was observed in diversity or richness, a shift was found in the EcM community composition supported by sporocarp, root tip and in-growth bag approaches (III). In addition, all taxa in humus samples that showed a response to lead were ectomycorrhizal fungi (II) and the differences followed the same trends in both datasets. Similarly to our results, Moser *et al.* (2005) analyzed EcM fungal communities by means of morphotyping and the pattern of restriction fragment length polymorphisms (RFLP), and reported that despite no observed difference in EcM fungal richness or diversity between the heavy metal-rich serpentine soils and control soils, there were differences in the community composition.

The total fungal community structure (II) and EcM fungal community structure (III) were altered under Pb contamination. The compositional shifts in the total fungal community were mainly attributable to the increased frequency of the genus *Thelephora* and to the decreased frequencies of the genera *Lactarius, Cortinarius* and *Piloderma* in the humus soil in contaminated areas compared to the control areas (II). In study III, we confirmed the increasing frequency of the genus *Thelephora* in in-growth bags. The taxa showing decreasing frequencies were different from II, namely the genera *Pseudotomentella, Suillus* and *Tylospora*, but *Lactarius, Cortinarius* and *Piloderma* did not show any difference (III). Based on these findings, it is difficult to draw conclusions on the lead sensitivity of EcM fungi. Moreover, lead-sensitive fungi may be able to colonize lead-free micro-environments in the polluted areas, as was earlier shown to be the case for bacteria (I).

*Thelephora terrestris* has been shown to exhibit tolerance to high lead concentrations in liquid cultures, although the author in this case suspected the precipitation of lead in the liquid media as a reason (Tam 1995). However, McCreight and Schroeder (1982) also reported the growth of *T. terrestris* on agar media only being arrested at a lead concentration of 200 µg ml\(^{-1}\). Together with our results (II, III), we can confirm that *T. terrestris* is lead tolerant. The genus *Thelephora* was relatively abundant in the contaminated area, while the whole family *Thelephoraceae* represented a very small
proportion of the total fungal diversity in the non-contaminated area (II). In paper III, both root tip and in-growth bag data revealed that *Thelephora* was identified most often in the Pb-contaminated area. The NMS test also demonstrated that the frequency of *Thelephora* was positively correlated with Pb enrichment on the Pb gradient from 37 to 49 700 mg Pb g\(^{-1}\) (total Pb concentration) in the humus layer (III).

*Cortinarius* did not show Pb sensitivity in in-growth bag data (III). However, an incongruence was observed in that although *Cortinarius* is the most species-rich genus in our sporocarp and root tip data, its frequency in in-growth bag data was relatively low (III). The genus *Cortinarius* has a relatively high diversity in boreal forests (Niskanen et al. 2008). Five *Cortinarius* OTUs were detected in the clean area while two were observed in the contaminated area (II), which suggests variation in lead tolerance among *Cortinarius* species. The higher abundance of *Cortinarius* in the clean area than in the contaminated areas may reflect lead sensitivity in some, but not all *Cortinarius* species.

The genus *Lactarius* was mainly composed of *L. rufus*, which has been reported as a common species in boreal coniferous forests (Väre et al. 1996, Perkiömäki and Fritze, 2005). Krupa and Kozdrój (2007) suggested that the ectomycorrhizal *L. rufus* is able to assist its host plant (pine seedling) by restricting the translocation of Zn, Cd or Pb from roots to shoots in a heavy metal contaminated area, indicating that this fungal species is one of the survivors under such contaminated circumstances. In this study, *L. rufus* was observed in the contaminated area, but at lower frequencies than in the clean area both in soil (II) and in in-growth bag data (III), although significant difference was only observed in the former case. These observations indicate that *L. rufus* may have partial lead resistance.

Some genera may differ significantly in frequency between clean and contaminated areas but nevertheless have little impact on the fungal community structure. For example, *Suillus* was significantly more frequent in clean than contaminated areas (III). Although *Suillus* generally tends to form large numbers of sporocarps and tends to be abundant in root tips in coniferous forests (Gardes and Bruns 1996, Dahlberg 2002), this was not the case for our study site. Although frequent among the sporocarps, it was infrequent among the root tips or in the in-growth bags (III) and did not appear to be frequent in soil samples (II).

### 4.4 Diversity of bacteria and bioavailability of lead

Earlier studies at Hävlälä have shown a reduction in bacterial biomass due to lead (Rantalainen et al. 2006, I), indicating a possible reduction in bacterial diversity due to lead. To characterize the bacterial communities, we cloned and sequenced a total of 1152 amplicons. After quality control, the final dataset contained 917 bacterial 16S rDNA sequences providing a sequencing depth of 153 ± 22 (mean ± 1SD) sequences per sample (II). Species richness (number of observed OTUs, Sobs), diversity (Simpson’s D and Shannon H) and evenness of the bacterial community were not significantly decreased in the polluted area as compare with the clean area (II). This
finding indicates that the bacterial community in the samples representing the contaminated area was not significantly different from that in the clean area. The extrapolated species richness ($S_{\text{Chao1}}$) exceeded the observed number of OTUs, but was not affected by lead pollution either (II). Similar results have been reported by Grandlic et al. (2006), who reported that the bacterial community structure of neither cultivable nor noncultivable species was affected by lead. Thus, the lack of significant changes in the diversity of the bacterial community disproved the hypothesis that the bacterial community structure varies under Pb contamination at Hälväla.

The low bioavailability of lead in contaminated areas (I) may explain the unchanged bacterial diversity. The toxicity of water extracts and soil slurries, as shown by bioavailability test 1 (I), was very low, indicated by the fact that constitutively luminescent reporter bacterium exhibited reductions in light output only when exposed to the litter samples of the NC area. The luminescence induction measurement (bioavailability test 2 in I), however, demonstrated that the local environment is not totally devoid of bioavailable lead, suggesting that micro-organisms are exposed to lead contamination, but at low level, even in the most contaminated areas. Secondly, the shooting range has been in use for decades, and the new sector is still actively used. This has given time for lead-resistant bacteria to adapt and enrich in the contaminated soil. Pointing to an adapted bacterial community is the fact that the frequency of lead-resistant cultivable bacteria was elevated in the lead contaminated soil (I). Bååth et al. (2005) also detected an increased tolerance of lead in the bacterial community of naturally lead-enriched forest soil. Last but not the least, although the proportion of lead-resistant cultivable bacteria had increased by two orders of magnitude in the contaminated Hälväla soil, the majority of the cultivable bacteria were still sensitive to lead, and did not grow on plates amended with water soluble lead (I). Since these lead sensitive bacteria were present at high numbers in the lead contaminated soil, this further suggests that the bioavailability of the lead is low.

### 4.5 Lead tolerance of bacteria

A critical strategy for bacteria to cope with high lead concentrations is to find micro-niches where lead does not form a selective pressure. Nevertheless, the proportion of cultivable lead-resistant bacteria colonies was only 0.14% of the total (I). This low proportion suggests that a difference in community structure between the clean and contaminated sites may be difficult to identify by characterizing the total community using molecular techniques. The result in paper II showed that no major variation in the bacterial community structure was detected when comparing the clean and contaminated areas. The structure of the bacterial community is illustrated at five taxonomic levels (from genus to phylum) in the pie diagrams for both areas (II). On a phylum level, Proteobacteria-related OTUs comprised approximately 41% of the bacterial population in both clean and polluted areas, and no statistically significant differences were observed between the areas in the frequency of the
major lower taxa within this phylum (α-proteobacteria and Rhizobiales). In addition, the frequency of OTUs affiliated with Acidobacteria (34%) and Actinobacteria (10.5%), and the lower taxonomic levels within these phyla, were not affected by lead contamination. The frequency of Planctomycetes in the clean area was twice than in the polluted one, but the difference was not statistically significant due to the low total number of these OTUs. The bacterial community observed at Hälvälä was similar to that observed in coniferous forest soil by Lesaulnier et al. (2008), especially at the higher taxonomic levels (phylum, class and order). The bacterial community structure in both the clean and polluted areas thus appeared typical for that type of environment. Some OTUs were found only in clean or polluted sites. However, the low number of sequences does not allow for definite conclusion to be drawn regarding the effect of lead contamination.

4.6 Representativeness of sampling and analysis techniques

The spatial structure in the populations of ectomycorrhizal fungi could be related to the functional morphology of a particular species (Agerer, 2001). For example, Cortinarius generally produces copious hyphal strands emanating from the mantle surface, which may help to exclude competitors (Pickles et al. 2010). Lactarius rufus, on the other hand, produces more diffuse external mycelia, possibly less suited to excluding competitors (Pickles et al. 2010). The above arguments indicate a possibility that spatial clustering may affect the results and lead to confounding co-correlates with insufficient sampling replicates. To minimize the bias affecting observed fungal communities, we always analyzed ten samples from ten locations in each study area to overcome this problem (I, II, III).

A variety of different approaches have been applied when studying microbial communities. It is extremely challenging to obtain a "perfect view" or even a 'good view' of the diversity and richness of these communities (Taylor 2002). For fungi, a sporocarp sampling survey is considered an inadequate method for investigation of the EcM community (Dahlberg 2001) due to the risk of bias against hidden or cryptic EcM fungal species. This notion is supported by the observations in III. The production of sporocarps is a poor indicator of the abundance of EcM species belowground (Dahlberg 2001, Taylor 2002). It is commonly accepted that the EcM community aboveground differs from that analyzed below-ground (Taylor 2002,
Root tip analysis can also only detect EcM fungal species that form mantles. However, our root tip results can be considered reliable, since among the top 10 genera of the 454 pyrosequencing dataset there was only one genus, *Wilcoxina* that does not form mantle structures (III).

DNA analysis may provide a more comprehensive picture of the microbial communities. We obtained 264 bacterial sequences from soil samples (I), 917 bacterial sequences and 627 fungal sequences in humus soil samples using DNA cloning and sequencing (II), and 8134 fungal sequences using 454 pyrosequencing in in-growth bags (III). If there is a positive relationship between sequencing effort and species recovery, high-throughput sequencing efforts that generate thousands of sequences per sample would presumably recover more species. However, the rarefaction analysis predicted unsaturated samplings even with the 454 technology (III). The purpose of the studies was, however, not to determine the total richness, but to assess whether lead contamination has had an influence on the community composition, and whether certain bacteria and fungi that are commonly found in boreal forest soil would be significantly reduced or totally eradicated by lead pollution. This goal was achieved.

### 4.7 454 pyrosequencing OTU assignment

When studying microbial diversity with a DNA sequencing method, a sequence similarity rate is needed for OTU assignment. Earlier, an ITS similarity of 90% or lower has been utilized by Kuninaga et al. (1997), while O'Brien et al. (2005) suggested that 97% or greater sequence similarity is probably a conservative measure of species richness, especially when sampling from a single locality, Buée et al. (2009) followed the later assumption and analyzed and revealed an unexpectedly high fungal diversity in forest soils. Recently, Jumpponen and Jones (2009) and Jumpponen et al. (2010) suggested 5% sequence variation as an OTU assignment cut-off value when studying fungal ITS2 sequences by 454 pyrosequencing. In order to test whether this value is suitable for our in-growth bag dataset which was produced by the same technique, the numbers of OTUs of EcM fungi were constructed under the 13 OTU assignment levels (Figure 3). The OTU richness was stable up to 95% sequence identity, but grew exponentially at thresholds greater than 95%, supporting the suggestion by Jumpponen and coworkers.

### 4.8 The new program for analyzing pyrosequencing data

MuSSA has not been utilized in the papers of this thesis, since the program had not been developed when preparing paper III, which contains 454 pyrosequencing data. However, the development of MuSSA has greatly profited from the experience of data handling for paper III. In the autumn of 2009, MuSSA was used by a project that concerns organic contamination in soil. Fifty soil samples were analyzed and the partial 16s rRNA gene of bacterial community was 454 pyrosequenced. MuSSA quality control system filtered out 3302 sequences that are shorter than 150 bp, without primer and ID tag sequence and with more than 0 chimerical nucleotide characters. Finally
altogether, 24038 sequences were assigned at five similarity levels (91, 93, 95, 97, 99%; Figure 4) and frequency matrixes were generated for each of the 5 similarities (not shown here due to the space required).

Besides MuSSA, DOTUR (Distance-Based OTU and Richness, Schloss et al. 2004) is another similar OTU assignment program. DOTUR assigns sequences to OTUs by using either the furthest, average, or nearest neighbor algorithm for each distance level. This program uses a PHYLIP (http://evolution.genetics.washington.edu/phylip.html) distance matrix as an input file, which assign sequences to OTUs for every possible distance, while MuSSA uses CAP3 (Huang and Madan 1999) in computation of overlaps between sequences, and construction of multiple sequence alignments of sequences. To compare the performances of the two programs in assigning sequences to OTUs, 498 bacterial partial 16s rRNA gene sequences, randomly selected from the dataset mentioned above, were tested by both programs. Results are as shown in Table 6.

DOTUR constructs randomized rarefaction and collector's curves of observed OTU, diversity indices, and richness estimators, while the current version (1.5) of MuSSA does not have these functions yet due to the original limited development budget. However, by slightly modifying output files of MuSSA into input files, EstimateS can easily carry out such functions. A later updated version of MuSSA will include this type of functions.

Figure 3. The numbers of operational taxonomic units (OTU) under the 13 OTU assignment levels in control, newly contaminated and old contaminated areas. Error bars indicate ± one standard deviation (1SD).
In 2009, DOTUR was integrated in MOTHUR, which is a package of popular DNA tools, including DOTUR, SONS, s-libshuff, TreeClimber (i.e. the parsimony test), UniFrac, distance calculation, visualization tools, a NAST-based aligner, and many other features. With the assistance of such functions as "collect.shared", "rarefaction.shared", and "dist.shared", MOTHUR is able to treat multiple samples for many novel analyses. MOTHUR is a command based program with a number of complicated and ambiguous operational steps, and it is difficult to manipulate such a program for ordinary people used to windows and simple operations. When doing sequence assignment, MuSSA is easy to learn as a result of windows based control panel which displays most parameters and function buttons. As a new developed product, localized tips were included in the program design, the input and output files are formulated according to the local workers' habits.

Figure 4. Following the analysis of 50 soil samples (n=50), a total of 24038 partial 16S rDNA sequences were included in the 454 pyrosequencing dataset. OTUs were calculated under 5 assignment levels by MuSSA. Error bars indicate ± one standard deviation (1SD). Unpublished data, with the permission of Aki Sinkkonen.
Table 6. Comparison of MuSSA and DOTUR

<table>
<thead>
<tr>
<th></th>
<th>MuSSA</th>
<th>DOTUR (integrated in MOTHUR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method based on Control panel</td>
<td>CAP3 Windows based</td>
<td>PHYLIP Command based</td>
</tr>
<tr>
<td>Support multiple sample analyses</td>
<td>Yes</td>
<td>Function can be achieved with the assistant of MOTHUR, but more complicated more time consuming than MuSSA.</td>
</tr>
<tr>
<td>Sequence quality control</td>
<td>Yes</td>
<td>Function can be achieved by other tools included in MOTHUR</td>
</tr>
<tr>
<td>Rarefaction and diversity indices</td>
<td>No(^1)</td>
<td>Yes</td>
</tr>
<tr>
<td>Test run with bacterial partial 16s rRNA gene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of sequences tested</td>
<td>498</td>
<td>498</td>
</tr>
<tr>
<td>No. of OTUs at 93% similarity</td>
<td>125</td>
<td>119(^2)</td>
</tr>
<tr>
<td>No. of OTUs at 95% similarity</td>
<td>157</td>
<td>147(^2)</td>
</tr>
<tr>
<td>No. of OTUs at 97% similarity</td>
<td>210</td>
<td>197(^2)</td>
</tr>
</tbody>
</table>

\(^1\)The version 1.5 does not contain this function yet, but with the assistant of EstimateS (Colwell et al. 2006), the functions can be achieved.

\(^2\)Using Furthest Neighbor Algorithm
5 CONCLUSIONS

The DNA cloning and sequencing approaches used in this study were revealed to be effective tools in studying the differences in microbial communities of the soil between lead contaminated and non-contaminated coniferous forest areas at Hälvälä shooting range. No effect of lead contamination was found on bacterial richness or diversity, while fungal richness and diversity were significantly reduced by lead. Some OTUs assigned to Basidiomycota were clearly affected and were rarer in the lead contaminated areas. However, even in the case of fungi, genera that were deemed sensitive were not totally absent from the contaminated area: only their relative frequency was significantly reduced. The observed difference between bacteria and fungi in their response to lead suggests a difference in their strategy to avoid the contaminant. Alternatively, bacterial species may adapt more readily by acquiring new traits, thereby avoiding being replaced by new species. The analyses performed here do not, however, allow for definite conclusions regarding these mechanisms.

454 pyrosequencing provided a detailed picture of the effect of lead on EcM fungal communities. Although organismal diversity and richness seemed unresponsive to long-term Pb contamination, the EcM fungal communities shifted in composition. This trend was visible in the sporocarp and root tip data and appeared statistically significant in the in-growth bag data. These compositional shifts were probably a result of the increase in abundance of *Thelephora*, and decreases in the abundances of *Pseudotomentella, Suillus* and *Tylospora* in contaminated areas relative to the control. The enrichment of *Thelephora* in contaminated area was also observed when studying the total fungal communities in soil using DNA cloning and sequencing technology. However, it remains unclear whether these community shifts have functional consequences for the dominant trees or soil ecosystem.

Organisms representing different trophic groups adapt to the presence of lead in different ways. While plants may respond by growth retardation, the bacterial microflora responds and adapts with a partial change in diversity, which in this thesis was only observed as an elevation in the frequency of lead-resistant cultivable bacteria. The results presented here point to a need for a versatile approach when assessing the environmental impact of lead contamination in forest soil. Each single test type, test organism and set of testing conditions only tells a small part of the whole story. New tools are needed to begin to approach the whole story, and the computer program Multi-Sample Sequence Analyser (MuSSA) was developed as a part of this thesis. The greatest advantage of the program is that large volumes of sequence data can be manipulated and general OTU series with frequency values can be calculated from a large number of samples.

In general, the results of the studies forming this thesis indicate that at the Hälvälä shooting range, lead influences the fungal communities but not the bacterial communities. The forest ecosystem shows apparent functional redundancy, since no significant effects were seen on forest trees.
This work was carried out at the Department of Environmental Sciences, University of Helsinki. Helsinki University Centre For Environment (HENVI), Finnish Cultural Foundation and Finnish Graduate School in Environmental Science and Technology (EnSTe) are acknowledged for the financial support.

I wish to express my warmest gratitude to my supervisors Professor Martin Romantschuk and Professor Rauni Strömmer. Mara, I thank you for the opportunity to study towards PhD under your supervision. In the past years, you provided me with many valuable advices during all the steps of my PhD and exhibited a real consideration on my thesis work. I always gain outstanding ideas from you when I am trapped in a puzzle. Rauni I thank you for always having time for comments and guidance. I appreciate your knowledge on fungal ecology and introducing of Dr. Ari Jumpponen. Ari, I want to thank you for teaching me how to work with 454 pyrosequencing in microbial ecology. I am also grateful to Professor Timo Kairesalo for providing me the opportunity to complete this work at the Department of Environmental Sciences at Lahti.

I wish to thank Senior Researcher Jaak Truu and Professor Marko Virta for reviewing the thesis and their comments for improving it. I also wish to thank my PhD study committee members Merja Kontro and Petri Auvinen for kind advices during my study. My co-authors, Juha Mikola, Tuula Niskanen, Jukka Kurola, Joanna Hanzel, Teija Koivula, Xinxin Liu, Salla Selonen and all other co-authors are kindly acknowledged for their expertise and smooth collaboration.

Egbert Beuker, I will never forget you since you are the person who brings me to the science world. I also would like to thank Martti Venäläinen and Tuija Aronen for being my supervisors when I worked in METLA Punkaharju Station.

Xinxin, my dear, thank you for considerate care at all stages of my studies and your contribution to the work. Mum and Dad, thanks for your encouragement and support on my study. Fenni, my lovely daughter, congratulations, you made your first step, so did I. I love you all forever!

Lahti, September 2010
7 REFERENCES


Benitez M.S. and Gardener B.B.M. 2009. Linking Sequence to Function in Soil Bacteria: Sequence-Directed Isolation of
Novel Bacteria Contributing to Soilborne Plant Disease Suppression. Applied and Environmental Microbiology 75, 915-924.


Costa R., Gotz M., Mrotzek N., Lottmann J., Berg G. and Smalla K. 2006. Effect of site and plant species on rhizosphere community structure as revealed by molecular analysis
of microbial guilds. FEMS Microbiology Ecology 56, 236-249.


Fitter A.H., Gilligan C.A., Hollingworth K., Kleczkowski A., Twyman R.M. and


Heylen, K., Vanparys B., Wittebolle L., Verstraete W., Boon N. and De Vos P. 2006. Cultivation of denitrifying bacteria: optimization of isolation conditions and


Krupa P. and Kozdrój J. 2007. Ectomycorrhizal fungi and associated bacteria provide protection against heavy metals in inoculated pine (Pinus sylvestris L.)


Environmental and Experimental Botany 22, 1-7.


