Lehrstuhl für Bodenökologie
der Technischen Universität München

Structural and functional diversity of proteolytic genes in an arable field

Mirna Mrkonjic Fuka

Vollständiger Abbdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende: Univ.-Prof. Dr. I. Kögel-Knabner

Prüfer der Dissertation:
1. Univ.-Prof. Dr. J. C. Munch
2. Univ.-Prof. Dr. S. Scherer

Die Dissertation wurde am 07.12.2006. bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 24.04.2007. angenommen.

I. Index

I. Index **I**

II. List of abbreviation **IV**

1. Introduction **1**

1.1 Role of organic nitrogen in soil 2

1.2 Origin of soil proteases 4

1.3 Classification of protelytic enzymes 5

1.4 Proteases as extracellular enzymes 6

1.5 Methods for the assessment of proteolytic microbial communities and their activity in soil 7

1.6 Hypothesis and objectives 10

2. Materials and methods **12**

2.1 Experimental site and sampling 12

2.2 Measurement of the abiotic soil parameters 13

2.3 Molecular analysis of microbial communities 14

2.3.1 Buffers and media 14

2.3.1.1 LB medium 14

2.3.1.2 NB medium 14

2.3.1.3 50 x TAE buffer 14

2.3.2 Nucleic acid extraction 15

2.3.2.1 Chromosomal DNA isolation from pure cultures 15

2.3.2.2 High molecular DNA isolation from the soil samples 15

2.3.2.3 Isolation of plasmid DNA (low yield) 16

2.3.2.4 Isolation of plasmid DNA (high yield) 16

2.3.3 Quantification of double-stranded DNA in solution 16

2.3.4 Standard PCR amplification 17

2.3.5 Agarose Gel Electrophoresis 18

2.3.6 Purification of PCR products 19

2.3.6.1 Purification of longer PCR products 19

2.3.6.2 Purification of high yield PCR products 19

2.3.6.3 Purification of short PCR products 19

2.3.7 Quantitative “Real-time” PCR 20
2.3.7.1 Amplification and cloning of *sub*, *npr* and 16S rRNA standard DNA 21
2.3.7.2 Real-Time PCR protease gene assay 21
2.3.7.3 TagMan 16S rRNA PCR assay 21
2.3.7.4 Examination of sensitivity of Real-Time PCR and DNA extraction recovery 22
2.3.8 Terminal restriction fragment length polymorphism (T-RFLP) 23
2.3.8.1 Selection of restriction enzymes 24
2.3.8.2 PCR amplification and purification of the PCR products 25
2.3.8.3 Restriction endonuclease digestion 25
2.3.8.4 Detection and analysis of *sub* and *npr* T–RFs 26
2.3.8.5 Analysis of T-RFLP data 27
2.3.9 Cloning analysis 28
2.3.9.1 Design of reverse *npr* PCR primer for phylogenetic analysis 28
2.3.9.2 Cloning 28
2.3.9.3 Sequencing 29
2.3.9.4 Purification of sequencing reaction 29
2.3.9.5 Sequence analysis 29
2.3.9.6 Nucleotide sequence accession numbers 30
2.3.9.7 Assignment of cloned *npr* sequences to T- RFLP 30
2.4 Microbial activity measurements 31
2.4.1 Proteolytic activity analysis 31
2.5 Statistical analysis 31

3. Results **33**
3.1 Abiotic soil parameters 33
3.2 Isolation of DNA from soil samples 37
3.3 Products of soil *npr* and *sub* PCR amplification 38
3.4 Quantification of *npr*, *sub* and 16S rRNA 39
3.4.1 Plasmid standard preparation and calculation of gene copies number 39
3.4.2 Examination of Real-Time PCR detection limits and DNA extraction efficiency 39
3.4.3 Abundance of *npr*, *sub* and 16S rRNA gene copies in soil samples 42

3.5 Proteolytic activity potential 44
3.6 Relation between soil bacteria protease genes, protease activity and 16S rRNA numbers 45
3.7 T-RFLP analysis 47
3.7.1 Optimization and specificity confirmation of T-RFLP analysis 47
3.7.2 Evaluation of soil T-RFLP assays 49
3.7.2.1 *Npr* T-RFLP evaluation 49
3.7.2.2 *Sub* T-RFLP evaluation 52
3.8 Establishment of *npr* gene bank 55
3.8.1 Rarefaction analysis 58
3.8.2 Specificity of the *npr* primer set used for phylogenetic analysis 59
3.8.3 Connection of cloned *npr* sequences to T- RFLP 60

4. Discussion 61
4.1 Usefulness of the primers used for analysis of total and proteolytic communities in soil samples 61
4.2 Efficiency of soil DNA extraction 62
4.3 Number and composition of 16S rRNA, *sub* and *npr* genes pro genome 63
4.4 Factors affecting abundance of 16S rRNA and protease genes in soil samples 64
4.5 Factors affecting potential proteolytic activity in soil samples 68
4.6 Comparison of *npr* and sub gene communities by T-RFLP 71
4.6.1 Comparison of *npr* and *sub* T-RFLP profiles in soil samples 72
4.7 Phylogenetic analysis of soil *npr* coding community 75
4.8 Conclusion and perspectives 78

5. Summary 79
6. References 80
7. Appendix 89
7.1 Figure legends 89
7.2 Table legends 92

8. Acknowledgments 93

II. List of Abbreviation

°C	degree centigrade
β	beta
bp	base pair
amp	ampicilin
BSA	bovine serum albumin
C	carbon
$CaCl_2$	calcium chloride
C_{org}	organic carbon
DGGE	denaturing gradient gel electrophoresis
dH2O	distilled water
DMSO	dimethyl sulfoxide
DNA	deoxyribose nucleic acid
DNase	deoxyribonuclease
EDTA	ethylene diamine tetra acetic acid
e.g.	for example
et al.	et alteri
g	gram
H	hydrogen
h	hours
ha	hectare
kg	kilogram
l	litre
LB	Luria Bertani (-medium)
µ	micron (10^{-6})
M	molar
m	milli (10^{-3})
Mg	magnesium
min	minute

mm	millimeter
MPN	most probable number
N	nitrogen
N_{tot}	total nitrogen
n	nano (10^{-9})
O	oxygen
PCR	polymerase chain reaction
pmol	pico moles (10^{-12})
RNA	ribose nucleic acid
rRNA	ribosomal RNA
SSCP	single strand conformation polymorphism
TAE	tris acetic acid EDTA buffer
T-RFs	terminal restriction fragments
T-RFLP	terminal restriction fragment length polymorphism
V	volt
vol.	volume
w/v	weight / volume
yr	year

1. Introduction

Nitrogen is an essential element for soil fertility and plant nutrition. It is a limited resource in soils, and N availability is one of the factors regulating organisms' growth in these ecosystems. The requirements of the plants and microorganisms for nitrogen are enormous, e.g. the uptake of nitrogen by wheat is about 85 kg/ha over one vegetation period as nitrogen is, next to carbon, the major nutritional element in plants. The usual agricultural management practice aims at maximizing productivity and at compensating N export by harvest. Therefore nitrogen is applied as fertilizer to non legume cropping systems and adds to organic N mobilized from organic substrate present in the soil. The risk of an excessive N application and contamination of ground and surface water by nitrate leaching can be avoided by a proper nitrogen management plan. Thereby, it is needed to fully understand the naturally occurring processes of nitrogen mobilization from the soil organic matter as organic nitrogen represents the largest and most important soil N pool. Since proteins are the major part of soils organic nitrogen compounds, their mineralization and degradation stands for the major process within the nitrogen cycle. The indigenous bacterial community harboring genes encoding for extracellular proteases plays a critical role in regulating proteolysis and nitrogen transformation in soils. Proteases mediate the conversion of unavailable forms of nitrogen into forms that are readily assimilated by plants or microbial biomass and serves as N but also C, H and O source. Hydrolysis of polypeptidic compounds is predominant for the global N-turnover because soils´ proteolytic activity influences the intensity and direction of biochemical processes of N transformation. As such activity of soil native proteolytic community is of prime importance in maintaining high productive agriculture in less environmentally damaging way. However, the response of the indigenous proteolytic community on environmental stress and perturbation is not well understood. A better knowledge of the size and structure of bacterial proteolytic genes community, extracellular proteases activity and environmental factors influencing microbial ability to function in the soil, is needed. The mechanism involved in microbial response to environmental variability and changes is of particular importance. Studies addressing these effects on soil proteolytic bacterial community are limited and are not sufficient. Therefore, the investigation of the properties of proteolytic bacterial community by targeting the genes involved in proteolysis is crucial. The aim of present study was to profound our knowledge about functional and structural aspects of soil native proteolytic community with focus on factors influencing its size, structure and activity under the set of naturally occurring environmental conditions.

1.1 Role of organic nitrogen in soil

Nitrogen occurs in soil in several different organic and inorganic forms. Organic nitrogen represents a significant pool of total nitrogen in terrestrial systems. The surface soil of most cultivated areas contains between 0.06 and 0.3% N of which over 90% occurs in organic forms. The dynamics of organic nitrogen is reflected by mineralization, immobilization, leaching and plant uptake and it is more constant than mineral pool size (Murphy *et al.*, 2000). The plough layer of arable soils may contain more than 3000 kg N ha^{-1} (Stevenson, 1982), and most of it is composed of complex organic forms that could be separated into a number of pools (Paul & Juma, 1981). These pools are composed of organic N that is virtually inert or N present in the living cells as soil microbial biomass (Harmsen & Kolenbrander, 1965). However, only a part of this organic nitrogen pool is available to plants and microorganisms. It is assumed that bioavailable portion of organic nitrogen comprises about 4-6% of total nitrogen in soil being mostly the part of microbial biomass (Murphy *et al.*, 2000). Jarvis *et al.* (1996) concluded that arable soils contain usually about 1% of mineral nitrogen from the total nitrogen amount, except after mineral fertilizer application, such be significantly lower than organic nitrogen content. As suggested by Schulten & Schnitzer (1998) the soil bioavailable organic nitrogen pool is composed as followed: 40% proteinaceous material, 35% heterocyclic compounds as nucleic acids, 5-6% amino sugars, and 19% ammonium.

Nitrogen enters the soil in organic forms such as plant materials, in addition to dead animals and microorganisms, manure, compost, and sewage sludge. Proteins represent the largest inputs of soil organic nitrogen (Lipson & Näsholm, 2001) and therefore, their ecological importance is predominant. Proteases decompose proteins into smaller membrane-permeable products, which living organisms can assimilate and metabolize (Loll & Bollag, 1983) thus transferring nitrogen into the N cycle (Figure 1).

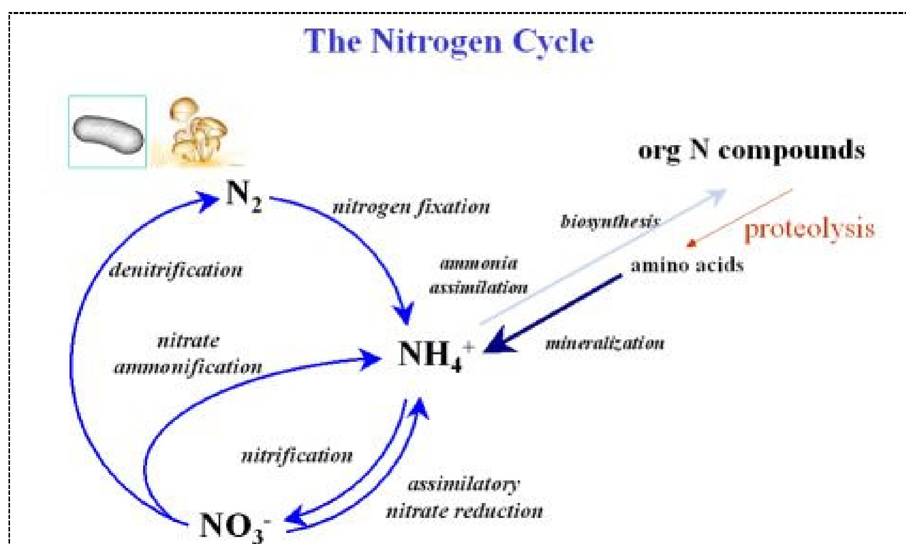


Figure 1: Nitrogen cycle (modified from K. Kloos, not published data)

Without the hydrolysis of proteins (proteolysis), nitrogen would be locked in a form that cannot be used by others (Ladd & Butler, 1972; Skujins, 1976). Thus extracellular proteases are primarily involved in maintaining continuum of life on earth by providing cells with N as amino acids or ammonium ions as proteolysis is generally considered to be rate-limiting step in N mineralization (Asmar *et al.*, 1994).

Mainly in soils without mineral fertilization, proteolysis is besides nitrogen fixation the only way how plants get access to nitrogen (Jones & Kielland, 2002). In low input farming and in systems with animal husbandry, proteolytic activity is a key process in the degradation of nitrous compounds, thus the soil ecosystems are supplied with easily accessible nitrogen for plants and microorganisms. Because of their beneficial role, proteolytic activity in soil is critical to N sustainability through providing a low cost and renewable supply of N that is less exposed to leaching and environmental pollution.

However bioavailability of proteins and amino acids is limited not only by the level of proteolytic activity but it is controlled by tendency of proteinaceuous material to adsorb to humic substances or soil minerals. As such, soil matrix acts as both as a sink and as reservoir for organic nitrogen. Although Rowell *et al.* (1973) observed significant activity of protease-humic acid complexes, Marshman & Marshall (1981) noticed decreased level of bacterial degradation caused by binding of proteins to the clay particles.

1.2 Origin of soil proteases

Proteases are ubiquitous and can be found in a wide diversity of taxa across the plants, animals and microorganisms. The origin and contribution of animals, plants and microbes proteases is discussed controversially in the literature. Ji and Brune (2005) reported about significant influence of soil-feeding termites in tropical rain forests and savannahs to the cycling of nitrogen due to the mobilization and digestion of soil peptidic compounds by a proteolytic activity of termites gut extracts. However in temperate systems the contribution of animal proteases to global N turnover is neglect able as those of plant origin. Although Badalucco *et al.* (1996) noticed high proteolytic activity in the root hair-zone of wheat plants and concluded therefore a significant influence of root-hair-enzymes to the overall protease activity, Hayano (1983) found much lower protease activity in tomato plants per root dry weight as compared to phosphatase and ß-glucosidase, respectively. Hayano (1996) concluded that the contribution of proteases secreted directly from plants is probably minor or much less significant than that of other soil enzymes. Additionally, it has been shown that microbial proteases have a higher substrate affinity compared to animal and plant proteases thus being the main source for N mineralization in soil (Ji & Brune, 2005).

Law (1980) pointed about significant role of extracellular proteolytic enzymes derived from various bacteria, although it is not always clear whether the enzymes described are truly extracellular or are released by dead cells. These findings are supported by studies of Hayano (1996) and Watanabe and Hayano (1994) that revealed no contribution of fungal protease and remarkable importance of bacterial proteases to soil proteolytic activity after selective inhibition of respective microbes. Bacterial proteolytic activity is especially crucial in agricultural soils due to the cultivation (tillage) and management (fungicide application) practice that is not fungal apt.

The indigenous bacterial community harboring genes encoding for extracellular proteases play a critical role in regulating proteolysis and nitrogen mobilization in soils. Numerous studies revealed differences in proteolytic activity by different soil physicochemical properties, climatic conditions and management practice (Marx *et al.*, 2005; Schloter *et al.*, 2003; Watanabe *et al.*, 2003). Although proteolytic activity has been intensively investigated during the last few decades, only limited research have been done to explore microbial sources of soil proteolytic enzymes and even less, influence of environmental or

anthropogenic factors on native proteolytic bacterial community. The early efforts made by Bach and Munch (2000) were based on the isolation and MPN count of bacteria with proteolytic potential. They investigated four topsoils and three subsoils, with different ecophysiological properties, for most abundant proteolytic bacteria. No coherence was observed between site-specific properties and MPN counts of proteolytic bacteria. In subsoil MPN counts were significantly lower only in the arable soil profile in March and no in October after harvest. Development of genetic markers and molecular tools has extremely afforded investigation of proteolytic bacteria in soil. Real-time PCR analysis did not show only a higher abundance of extracellular proteases genes in high yield than in low yield sites, but also clear effect based on the used farming management system (conventional or precision) (Bach *et al.*, 2002).

1.3 Classification of proteolytic enzymes

Classification of proteolytic enzymes is based on three major criteria: (i) type of reaction catalyzed, (ii) chemical nature of catalytic site, and (iii) evolutionary relationship with reference to structure (Rao *et al.*, 1998). In general, proteases are divided into two major groups (i) endoproteases that act on the interior linkages of substrate and (ii) exoproteases that attack free ends of proteins. Microbial proteases are predominantly extracellular since a direct uptake of proteins does not occur as such (Law, 1980; Loll & Bollag, 1983). They are separated into four prime groups based on the functional group present at the active site: serine proteases (EC 3.4.21), aspartic proteases (EC 3.4.23), cysteine proteases (EC 3.4.21), and metalloproteases (EC 3.4.24). Proteases are further divided in different families and clans depending on their amino acid sequences and evolutionary relationships. Based on the pH of their optimal activity, they are classified as acidic, neutral, or alkaline proteases. Predominant fungal extracellular proteases are supposedly cysteine and aspartic proteases, whereas those of bacterial origin are mainly alkaline metalloproteases (Apr), neutral metalloproteases (Npr) and serine protease (Ser) (Kalisz, 1988). Selective inhibition of protease activity indicated that the most common arable soil proteases belong to the neutral metalloproteases (Bach & Munch, 2000; Hayano *et al.*, 1987; Hayano *et al.*, 1995; Kamimura & Hayano, 2000) or are the part of the serine proteases group (Kamimura & Hayano, 2000; Watanabe *et al.*, 2003).

Metalloproteases are characterized by the requirement for a divalent ion for their activity. The neutral metalloproteases show specificity for hydrophobic amino acids whereas the alkaline

metalloproteases possess broad substrate specificity. The characteristic of metalloproteases is the irreversible inhibition of their activity by chelating agents such as EDTA. The genes encoding for alkaline (*apr*) and neutral (*npr*) metalloproteases are found in Gram-positive and Gram-negative bacteria.

Serine proteases are characterized by the presence of a serine group in their active site and exhibit wide substrate specificity. Serine proteases are recognized by their irreversible inhibition by 3,4-dichloroisocoumarin (3,4-DCI), $_{L}$-3-carboxytrans-2,3-epoxypropyl-leucylamido (4-guanidine) butane (E.64), diisopropylfluorophosphate (DFP), phenylmethylsulfonyl fluoride (PMSF), and tosyl-$_{L}$-lysine chloromethyl ketone (TLCK). Serine alkaline protease are widespread among viruses, bacteria, or eukaryotes and subtilisin, as a second largest family of serine proteases is until now only found among Gram-positive *Bacillus* ssp. Gene encoding for subtilisin (*sub*) may also serve as pathogenic factor as is concluded in study by Qiuhong *et al.* (2006) that investigated protease from nematode killer *Bacillus* sp. B16, showing the 98% similarity of deduced amino acids to the subtilisin from other *Bacillus* species.

1.4 Proteases as extracellular enzymes

The extracellular bacterial proteases are synthesized as inactive prepro-enzymes with an additional polypeptide segment (propeptide) that is removed from the mature (secreted) protein (Schmidt *et al.*, 1995; Takagi *et al.*, 1985; Wandersman, 1989). The biological importance of propeptide is still under discussion. One of the possible roles of the propeptide is to keep protease inactive insight the cell, thus protecting the cell against the proteolysis; it may play an important function in the folding of prepro-enzymes into the proper conformation necessary for the secretion or activity; and it might have function in temporal stabilization of proteases to the cell membrane. The cleavage mechanism to the mature form is still not known although the autoproteolytic processing is the most suggested one.

Proteases are mostly secreted by a sec-dependent general secretion pathway as for many extracellular enzymes (Kim & Kendall, 2000). Regulation of protease production has been manly investigated at the physiological and nutritional level. However, in several organisms

regulatory proteins involved in protease gene transcription were identified (Haese & Finkelstein, 1993). Among several environmental factors temperature, pH as well addition of high energy substrate are known to be involved in protease production (Rahman *et al.*, 2005).

1.5 Methods for the assessment of proteolytic microbial communities and their activity in soil

Culture-dependent methods to study organisms involved in particular process are based on isolation of soil microorganisms on special selective media. To estimate the number of proteolytic ogranisms, plate counting and most probable number (MPN) techniques have been applied (Bach & Munch, 2000). Traditional methods to determine the number and composition of the bacterial communities require various culture techniques that are time-consuming and laborious. Additionally, alternation of original environmental parameters during cultivation can change the community structure through the application of new selective conditions. In effect a new community structure does not represent the *in situ* one. The proportion of organisms that can be cultured is only a small part of the total population (Amann *et al.*, 1995). Use of culture-dependent techniques comprises less than 1% of micro biota present in the soil. According to Sorensen (1997) this increases to 10% in rhizosphere samples. Therefore, our understanding of bacterial dynamics has been restricted due the biases introduced by traditional culture dependent methods (Sorensen, 1997).

The application of molecular biological tools has greatly afforded the study of microbial communities in natural environments. Genomic approach based on the isolation of DNA/RNA without the need to isolate or culture a species, allow for the better insight into the bacterial community properties. The analysis of 16S rRNA as universal phylogenetical marker has increased our knowledge about physiology and function of bacteria in natural habitat (Liesack *et al.*, 1997) as well has greatly enhanced our ability to investigate bacteria in nature (Pace *et al.*, 1986). However the proteolytic bacteria are distributed over many different phylogenetic groups. This makes the use of 16S rDNA, as a marker for proteolytic bateria inappropriate.

The application of functional genetic markers has been of prime importance in the investigation of particular community or process in the terrestrial ecosystem. Functional

markers are structural genes that encode for a key enzyme such revealing the information about physiological feature of one so-called functional group. The establishment and application of functional macromolecules has been increased such numerous different structural genes have been successfully detected in various habitats. However investigation of abundance and diversity of genes of interest was not possible without development of molecular techniques to study particular genes.

PCR primers and probes for the detection of genes encoding for extracellular bacterial proteases were first designed by Bach et al. (2001). To estimate the number of genes in environmental samples PCR based quantitative techniques as MPN-PCR (Fredslund et al., 2001; Mantynen et al., 1997) and Real-Time PCR (Bach et al., 2002; Henry et al., 2004; Rousselon et al., 2004) have been employed. The MPN-PCR method focuses on the quantification of gene fragments at the end of amplification and requires post-PCR steps to ensure visualization of the generated products. Real–Time PCR techniques allow detection of the product once linear amplification is achieved and do not require post-PCR procedure. These techniques are more rapid, accurate and sensitive than MPN-PCR.

To study microbial diversity, numerous fingerprinting approaches have been developed (Kirk et al., 2004). However the most promising method to fingerprint proteolytic communities from the soil samples is the terminal restriction fragment length polymorphism technique (T-RFLP) so far. T-RFLP analysis is based on the restriction endonuclease digestion of fluorescently end-labeled PCR products. The digested products are separated by capillary gel electrophoresis and detected on an automated sequence analyzer. The technique provides diverse profiles (fingerprints) dependent on the species composition of the communities of the samples. T-RFLP analysis is successfully applied to asses a diversity of structural and functional genes as 16S rRNA, mercury resistance (*mer*), nitrite reductase (*nir*), ammonia monooxygenase (*amoA*), nitrate reductase (*narG*), nitrous oxide reductase (*nosZ*), and nitrogenase reductase (*nifH*) (Avrahami et al., 2002; Bruce, 1997; Delorme et al., 2003; Fedi et al., ; Horz et al., 2000; Rich & Myrold, 2004; Widmer et al., 1999; Wolsing & Prieme, 2004). Other commonly used methods to monitor highly diverse soil microbial communities are DGGE and SSCP (Schwieger & Tebbe, 1998; Sharma et al., 2005).

To gain detailed information about the composition of the proteolytic microbes, PCR products have to be cloned and sequenced. However, not necessarily the fragments used for quantitative PCR (qPCR) or fingerprinting are well suited for phylogenetic analysis as the length of the generated sequences is too short. Furthermore this approach is not suitable for high throughput analysis. At least 50 clones have to be analyzed per sample to get information on the most abundant species.

The presence of the target gene in complex communities can be simply demonstrated by hybridization of the DNA recovered from soil to the specific labeled probe. Early molecular studies used gene probe technology to screen for the presence or absence of structural genes in a soil population (Griffiths *et al.*, 2000). The technique has been successfully used to detect subtilisin (*sub*), neutral metalloprotease (*npr*) and alkaline metalloprotease (*apr*) gene fragments from the soil samples (Bach *et al.*, 2001) but also to detect transcripts of *sub* and *npr* genes in the rhizospheres (Sharma *et al.*, 2004).

Metagenomics represent a new approach in a genomic analysis. This method accesses the potential reservoir of novel genes in soil. To explore this reservoir, DNA from an environmental sample is extracted, cloned into an appropriate vector and transformed into competent *E. coli* cells. The resulting transformants in metagenomic libraries are screened for novel physiological, metabolic and genetic features. Although time-consuming and labour–intensive, metagenomic is the most powerful environmental approach that offers possibilities to discover novel genes and novel bio molecules through the expression of genes from uncultivated and unknown bacteria in recipient host cell. Theoretically, a metagenomic database should contain DNA sequences for all the genes in the microbial community (Handelsman, 2004). However, often those genes can not be expressed, folded or excreted correctly in the corresponding host system.

Soil proteolytic activity is studied indirectly by measuring the activity via assays since it is difficult to extract enzymes from soils (Dick, 1992). The assays are generally sensitive, short-term and reproducible. A small quantity of soil is incubated for a short period using dye- or fluorescence-labeled proteins as a substrate analogue. Proteolytic activity is usually detected by either: (i) following the decrease of initial substrate or more often (ii) measuring the increase of amino acids or peptides released during the incubation period. Detection of

released chromogenic or fluorogenic compounds can be measured spectrophotometrically or fluorometrically, respectively (Hoppe et al., 1988; Ladd & Butler, 1972). However, in vitro measurements of enzyme activities are rather showing a potential and do not reflect the in situ activity. During the measurements, conditions such as pH, temperature and substrate concentration are adapted to disclose the optimal activity that is often completely different from those of native systems. Furthermore, the proteolytic activity measurements do safer from the same limitation as assays used to analyze activity of others soil enzymes. However it reflects the proteolytic potential (Loll & Bollag, 1983) and indicates the biological capacity of a soil to carry out the process of proteolysis that is important for maintaining the soil fertility as soil fertility depends not only on nutrient status and availability but also on the nutrients turnover (Lopez-Hernandez et al., 1989). Additionally it gives us the opportunity to compare proteolytic potential of different soil samples or the samples that are treated differently.

1.6 Hypothesis and objectives

The indigenous bacterial communities harboring genes encoding for extracellular proteases play a key role in regulating proteolysis and nitrogen flux in soils. Thus there is a major interest in exploring factors influencing their size and activity. Additionally once secreted extracellular enzymes are no longer under the control of the bacteria. Therefore physical and chemical conditions of the habitat might exhibit a direct influence on the dynamics of extracellular protease activity. Therefore, the present thesis tent to investigate the soil proteolytic community in order to improve our understanding of proteolytic processes in soil as well to investigate the environmental factors affecting the ecology of bacterial community harboring genes encoding for extracellular proteases, mainly neutral metalloprotease (*npr)* and subtilisin (*sub*). In general, little is known about the functioning of bacterial proteolytic communities in soil. Therefore, more knowledge about the role of bacterial proteolytic community in regulating nitrogen mobilization from organic compounds is essential to understand the process and to facilitate the development of optimized management strategies.

The main hypothesis for this thesis is that bacterial proteolytic communities and their activity is highly dynamic in the soil system. It would exhibit high vertical, spatial and temporal variability due to the different physicochemical and nutrient status of different soil sites as

well as to the environmental changes that naturally occur at the study site during the season. To test this hypotheis we had the following aims:

(i) to evaluate the application of different molecular tools to quantify genes involved in protein degradation as well as to analyze the structure of *npr* or *sub* coding bacterial proteolytic community,

(ii) to characterized and to compare the proteolytic bacterial community at uniformly managed arable field with four different soil types and at three different depths,

(iii) to identify the effects of a set of naturally occurring environmental conditions with respect to abundance, structure, and activity of bacterial proteolytic community,

(iv) to investigate if there is a general relation between soil bacteria protease genes, protease activity and 16S rRNA numbers.

2. MATERIALS AND METHODS

2.1 Experimental site and sampling

The experimental site (A15) is located on the research farm "Klostergut Scheyern" in southern Germany, 40 km north of Munich (N48°30.0`; E11°20.7`). During the 1990`s as a part of the research program the conventional farming management was replaced by an integrated management, including minimum tillage and continuous crop cover. The climate is characterized by a mean annual temperature of 7.4° C and an average annual precipitation of 833 mm yr^{-1} (Auerswald & Kainz, 1990; Schroeder *et al.*, 2002). Soil samples were taken from the field A15 (4.9 ha) that shows large soil heterogeneity with four different soil types that are classified as: Gleyic Cambisol (GC); Luvisol/ Cambisol (LC); Cambisol/ Pseudogley (CP); and Gley/ Kolluvisol (GK) (Figure 2).

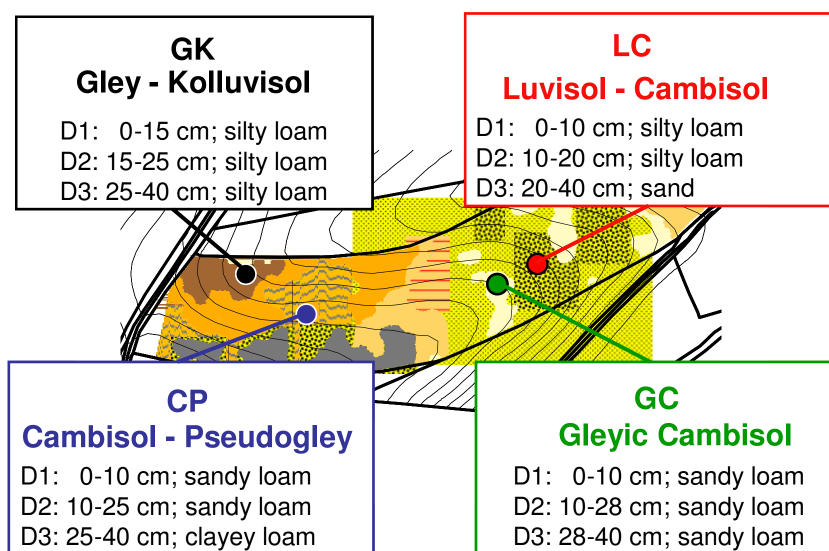


Figure 2: Heterogeneity and the main soil types of the field site (A 15) investigated in the present study

During the investigation the field was cropped with potatoes (*Solanum tuberosum* L.). Before cultivation of potatoes, winter wheat (*Triticum aestivum* L.) had been grown at the same experimental field. The field was managed according to good agricultural practice (Table 1). Bulk soil samples were collected using a soil auger (3 cm diameter). For each soil type, five soil cores were taken in close vicinity to each other. They were pooled and sieved (2 mm). The soil samples used for molecular based investigations were stored at -20°C and at 4°C for enzymatic approach.

Bulk soil samples were collected in 2003 on April 29th (before vegetation period), July 16th (development stage of the potato EC60) and October 20th (after harvest) from sites representing the four soil types at three different depths: D1 (≈0-10 cm), D2 (≈10-30 cm) and D3 (≈30-40 cm) according to the horizontation of the soil profile.

Table 1: Management of the field A15 in Scheyern in 2003

Date	Treatment/ Soil cultivation	Compounds	Aplication/ Soil depth
25.03.2003	Herbicide	Roundup	3.5 l ha^{-1}
	Fertilizer	UAN (Urea Ammonium Nitrate)	256 l ha^{-1}
25.04.2003	Rotary harrow and potatoes planting		10 cm depth
	Formation of the potatoes ridges		8 cm depth
29.04.2003	Formation of the potatoes ridges		8 cm depth
06.05.2003	Herbicide	Boxer	5 l ha^{-1}
	Herbicide	Sencor	0.2 ha^{-1}
11.06.2003	Fertilizer	Magnesium sulphate	25 kg ha^{-1}
	Fungicide	Ridomil	2 kg ha^{-1}
13.06.2003	Fertilizer	Urea	15 kg ha^{-1}
16.06.2003	Insecticide	Tamaron	1.2 l ha^{-1}
25.06.2003	Fungicide	Acrobat	2 kg ha^{-1}
	Insecticide	Karate	75 ml ha^{-1}
08.07.2003	Fertilizer	Magnesium sulphate	8 kg ha^{-1}
	Fungicide	Tanos	0 .7 kg ha^{-1}
25.07.2003	Fungicide	Mannex	2 l ha^{-1}
	Fungicide	Shirlan	0.2 l ha^{-1}
20.08.2003	Chemical desiccation	Reglone	2 l ha^{-1}
25.09.2003	Harvest of potatoes		23500 kg ha^{-1}

2.2 Measurement of the abiotic soil parameters

The dry matter content of the soil samples was determined gravimetrically by drying at 105°C for 24 h, and the pH was measured in 0.01 M $CaCl_2$ suspension according to Schlichting *et al.* (1995). Total organic carbon (C_{org}) and total nitrogen (N_{tot}) of soil samples were measured by a C/N analyzer (Carlo Erba 1500, Milano, Italy) with high catalytic combustion. Sieved soil samples were air-dried for 4 weeks before analysis on water extractable organic C and N (WEOC and WEON). Water extractable carbon and nitrogen was extracted with 10 mM $CaCl_2$ in an overhead shaker for 10 min at room temperature with a soil: solution ratio of 1:2 (w/v). After centrifugation (10 min, 3000 g), the supernatant was passed through a 0.45 µm

polycarbonate filter. Non-organic carbon was removed by acidification and purging the samples with pure O_2 for 2 min prior to measurement. Filtered solutions were quantified for WEOC using catalytic high temperature combustion (680°C) with a Total Carbon Analyser (Shimadzu TOC 5050A, Kyoto, Japan). The WEOC concentrations were referred to weighted soil mass (dry matter) and expressed as µg C g^{-1} dry matter. Dissolved total N, NO_3-N, and NH_4-N of the extracts were quantified photometrically with an automated continuous flow analyser (Skalar, Erkelenz, Germany). WEON was calculated by subtraction of mineral N (NO_3, NH_4) from total N.

2.3 Molecular analysis of microbial communities

2.3.1 Buffers and media

2.3.1.1 LB medium (Luria Bertani medium)

Peptone	10.00 g
NaCl	10.00 g
Yeast extracts	5.00 g

Volume was adjusted to 1 l using distilled water. pH was adjusted to 7.5.

2.3.1.2 NB medium (nutrient broth medium)

Peptone from meat	10.00 g
Meat extracts	3.00 g

Volume was adjusted to 1 l using distilled water. pH was adjusted to 7.0.

2.3.1.3 50 x TAE buffer

Tris base	242.00 g
EDTA	18.60 g
Glacial Acetic Acid	57.10 ml

Volume was adjusted to 1 l using distilled water. pH was adjusted to 8.0.

2.3.2 Nucleic acid extraction

2.3.2.1 Chromosomal DNA isolation from pure cultures

Bacterial strains listed in Table 2 were grown under aerobic conditions as recommended by German Microorganism Collection to an optical density <1 that was measured at 600 nm by spectrophotometer (CE 3021, Cecil Instruments Ltd., Cambridge, England). DNA was extracted using DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) as described by Kloos *et al.* (1998).

Table 2: Bacteria used for microbiological analysis

species	strain	cultivation condition (medium, temperature)
Bacillus cereus	DSM 3101	NB , 30°C
Bacillus subtilis	DSM 10	NB , 30°C
Bacillus thuringiensis	DSM 2046	NB , 30°C
Bacillus amyloliquefaciens	DSM 7	NB , 30°C
Bacillus licheniformis	DSM 13	NB , 37°C
Bacillus megaterium	DSM 90	NB , 30°C
Bacillus stearothermophilus	DSM 22	NB , 50°C
Bacillus sp.	DSM 405	NB , 70°C
E. coli	DSM 30083	LB , 37°C

DSM, German Microorganism Collection

2.3.2.2 High molecular DNA isolation from the soil samples

Soil DNA was extracted and purified using the FastDNA SPIN Kit for Soil (Bio 101, Vista, USA). 500 mg of soil sample was weight in “MULTIMIX 2 Matrix Tubes” and three independent DNA extractions were made as recommended by the manufacturer. Extracted DNA was immediately used or was stored at - 20°C. The quality of DNA was evaluated in 0.9% agarose gel in 1 x TAE buffer followed by 20 min staining with ethidium bromide (see section 2.3.5).

2.3.2.3 Isolation of plasmid DNA (low yield)

Plasmid DNA from the transformed *E. coli* competent cells was co-extracted and purified using NucleoSpin Plasmid Kit (Machery & Nagel GmbH & Co. KG, Düren, Germany) following the manufacturer's instructions. 3 ml overnight cultures (LB medium, 50 µg ml^{-1} of kanamycin) were harvested at 1100 rpm for 1 min. Pure plasmid DNA was finally eluted under low ionic strength conditions with 50 µl slightly alkaline buffer AE (5 mM Tris-Cl, pH 8.5).

2.3.2.4 Isolation of plasmid DNA (high yield)

To isolate plasmid DNA with high yields so-called "Midi preparation" was made using Qiagen Plasmid Midi Kit (Qiagen, Hilden, Germany). The procedure was done as recommended by the manufacturer. The starter culture of 3 ml of transformed *E. coli* (LB medium, 100 µg ml^{-1} of ampicillin) was shaked at 37°C and 250 rpm. The starter culture was diluted 1/5000 in LB *amp* medium. 50 ml of diluted culture was incubated at 37°C and 250 rpm for 14 hours. The bacterial cells were harvested by centrifugation at 6000 x g for 15 min at 4°C and pellets were further used for plasmid extraction. Ultrapure plasmid DNA was finally eluted with 150 µl TE buffer (10 mM Tris-Cl, pH 7.5, 1 mM EDTA).

2.3.3 Quantification of double-stranded DNA in solution

The amount of isolated DNA was estimated using the Quant-iT™ PicoGreen ® dsDNA Kit (Molecular Probes Inc., Eugene, OR, USA) according to the manufacturer protocol. 100 µl of blank probe, 100 µl of DNA standard dilutions and 100 µl of unknown samples were pipetted separately to the black 96-well strip plate. 100 µl of the diluted PicoGreen reagent was added to each probe and after 5 min of incubation measurement was done by fluorescence spectrophotometer (Cary Eclipse; Varian; Victoria, Australia). The wave lengths of the filters used were for excitation 485 nm and for emission 538 nm. From the subtracted average value from standards DNA the standard curve was generated allowing estimation of DNA concentration of the unknown samples.

2.3.4 Standard PCR amplification

Amplification was performed for different targets as describes in the Table 3. The composition of oligonucleotides used as primers are listed in Table 4. The volume (in µl) of different components used in PCR mix in 50 µl reactions is shown in Table 5. The primer stocks used were 10 pmol µl $^{-1}$ for 16S rRNA and 100 pmol µl $^{-1}$ for protease assays.

Table 3: Primers and cycling conditions used in this study

Target	Primers set (Reference)	Cycling Programme (denaturation / annealing / elongation)	No. of cycles	Expected fragment size (bp)
16S rRNA	FP 16S rRNA & RP 16S rRNA (Bach *et al.*, 2002)	94°C-1 min / 54°C-1 min / 72°C-1 min	35	263
subtilisin (*sub*)	*subIa & subII* (Bach *et al.*, 2001)	94°C-30 sec / 55°C-30 sec / 72°C-30 sec	35	319
Neutral metalloprotease A (*nprA*)	*nprI & nprIIa* (Bach *et al.*, 2001)	94°C-30 sec / 55°C-30 sec / 72°C-30 sec	35	233
Neutral metalloprotease B (*nprB*)	*nprI & nprIIb (*Bach *et al.* (2001) and this study)	94°C-1 min / 46°C-1 min* / 72°C-1 min	30	483

* Touch down PCR, annealing temperature was gradually decreased from 60°C for 2°C each two PCR cycles: Finally annealing was performed at 46°C as described above.

Table 4: Characteristic of oligonucleotides used as primers

Oligonucleotide	Composition	Position (nt)[a]
FP sub I	5´-ATGSAYRTTRYYAAYATGAG-3´	853- 872
RP sub II	5´-GWGWHGCCATNGAYGTWC-3´	1154- 1171
FP npr I	5´-GTDGAYGCHCAYTAYTAYGC-3´	214- 233
RP npr IIa	5´-ACMGCATGBGTYADYTCATG-3´	437- 446
RP npr IIb	5´-RTGDACNCCDCCRYWRT-3´	1427-1443
FP 16S rRNA	5´-GGTAGTCYAYGCMSTAAACG-3´	799- 818
RP 16S rRNA	5´-GACARCCATGCASCACCTG-3´	1044- 1063

[a] Nucleotide position in the *npr* gene of *B. cereus* (NCBI DQ129688), in the *sub* gene of *B. subtilis* (NCBI S51909) and in the 16S rRNA gene of *E. coli* (Brosius *et al.*, 1981). The protease targeting oligonucleotides are described in detail in Bach *et al.* (2001)

Table 5: Amounts of different components used in the amplification reaction

Component \ Target	16S	nprA	sub	nprB
Buffer (10x)	5	5	5	5
MgCl2 (25 mM)	-	3	6	3
dNTP (2 mM)	5	2	2	2
Primer (Forward)	1	0.5	0.75	0.75
Primer (Reverse)	1	0.5	0.75	0.75
Bovine serum albumin (3%)	5	5	5	5
Dimethyl sulfoxide (DMSO)	2.5	-	-	2.5
Template (DNA)c	1	2	2	2
Nuclease free water	27.5	30	26.5	27
RedGoldStar DNA polymerase	2	2	2	2

Prior to the cycling, a hot start of 94°C for 5 min was performed. RedGoldStar DNA polymerase (Eurogentec, Seraing, Belgium) was diluted in 1x Buffer to 2U and added to the reaction mix during the pause. Cycles were followed by a final extension at 72°C for 10 min. All PCR products were stored at 4°C for further analysis or were immediately used.

2.3.5 Agarose Gel Electrophoresis

PCR reactions (see section 2.3.4) were checked for products on 1.5% agarose gels (Biozym, Oldendorf, Germany) prepared in 1x TAE buffer (see section 2.3.1.3) and separated at 100 V for 1.5 h in 1x TAE buffer followed by 20 min staining with ethidium bromide (0.5 mg l^{-1}).

2.3.6. Purification of PCR products

2.3.6.1 Purification of longer PCR products

Purification of PCR products longer than 100 bp was performed using Qiaquick PCR Purification Kit (Qiagen, Hilden, Germany) as recommended by the manufacturer. To elute pure DNA 50 µl of EB buffer (10 mM Tris-Cl, pH 8.5) was used.

2.3.6.2 Purification of high yield PCR products

When was necessary to obtain very small elution volumes, MinElute PCR Purification Kit (Qiagen, Hilden, Germany) was applied. It allows recovery of extremely high concentrated DNA. Procedure was done as was described in Kit supplied handbook. To elute pure DNA 10 µl of EB buffer (10 mM Tris-Cl, pH 8.5) was used.

2.3.6.3 Purification of short PCR products

Purification of PCR products shorter than 100 bp (>40 bp) was completed using Qiaquick Nucleotide Removal Kit (Qiagen, Hilden, Germany) following manufacturer's instructions. Pure DNA was eluted in 30 µl of EB buffer (10 mM Tris-Cl, pH 8.5).

2.3.7 Quantitative "Real-time" PCR

Quantitative "Real time" PCR is a method used to quantify the genes during the progress of a PCR amplification process. It is based on the detection and quantification of the fluorescence produced by a fluorescent reporter. These fluorescent reporter molecules include dyes that bind to the double-stranded DNA (i.e. SybrGreen) or sequence specific probes (i.e. TaqMan probes). The fluorescence emitted by the reporter molecule increases due to the accumulation of the PCR product in a reaction with each cycle of amplification. During the exponential phase in Real-time PCR assay a fluorescence signal threshold is determined. A fixed fluorescence threshold is set significantly above the baseline measured during the first 15 cycles. The number of PCR cycles needful to emit enough fluorescent signal to reach this threshold is determined as the threshold cycle, or Ct value. These Ct values are directly proportional to the number of the target sequence in the assay and are a base for calculation of DNA copy number. Diagram of the Real-Time PCR amplification plot is shown in Figure 3.

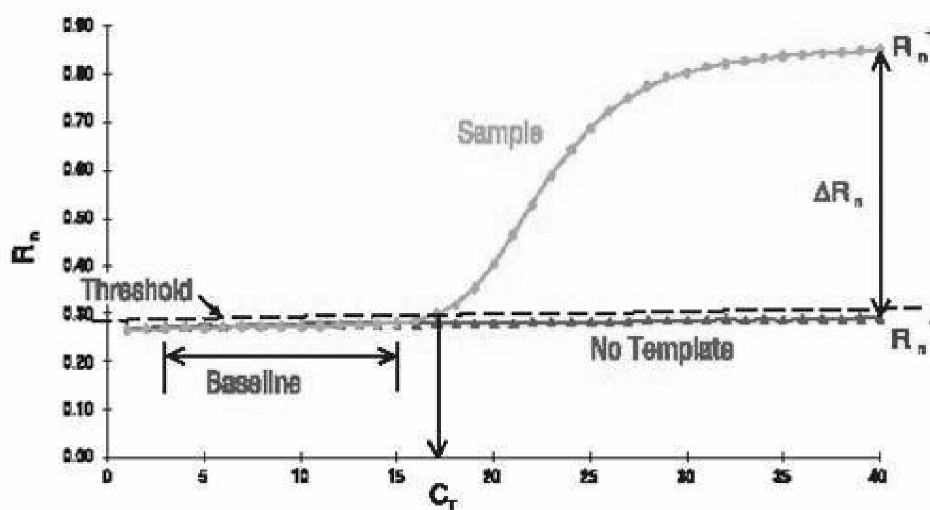


Figure 3: Diagram of the Real-Time PCR amplification plot (http://dna-9.int-med.iouwa.edu/realtime.htm)

Real-Time PCR approach was used to amplify gene fragments of 16S rRNA, serine proteases (*sub*) and neutral metalloproteases (*npr*). Amplification and monitoring was carried out on the ABI 7700 Sequence Detection System (Perkin Elmer, Norwalk, CT, USA) using SybrGreen as detection system when protease gene fragments were amplified and TaqMan probe as detection system during the amplification of 16S rRNA gene fragments. ABI 7700 Sequence Detection System software was applied for the quantification of the respective genes.

2.3.7.1 Amplification and cloning of *sub*, *npr* and 16S rRNA standard DNA

For the preparation of plasmid standards 16S rRNA, *sub* and *npr* gene fragments DNA from pure cultures was amplified by PCR using primers and conditions described in section 2.3.4. The obtained PCR products were purified using Quiagen PCR Purification Kit (see section 2.3.6.1), cloned into the pDrive Cloning Vector (Qiagen, Hilden, Germany) and transformed into QIAGEN EZ Competent Cells (Qiagen, Hilden, Germany) as recommended by manufacturer. The origins of the cloned sequences were as followed: *sub* (*Bacillus subtilis*), *npr* (*Bacillus cereus*) and 16S rRNA (*E. coli*) (see Table 2). Qiagen Plasmid Midi Kit (Qiagen, Hilden, Germany) was used to purify plasmid DNA (see section 2.3.2.4). Obtained plasmid DNA inserts were sequenced to confirm the origin (see section 2.3.9.3). Quantification of the plasmid DNA was done by PicoGreen assay (see section 2.3.3). Plasmid DNA was stored at – 20° C in 5 µl aliquots until used. The standard curves were created using tenfold dilution series of three plasmids containing *sub*, *npr* or 16S rRNA fragments.

2.3.7.2 Real-Time PCR protease gene assay

Amplification reactions for the protease gene assay were performed in volumes of 50 µl containing 25 µl of Platinum Sybr Green qPCR SuperMix- UDG including SYBR® Green I, 60 U ml^{-1} Platinum® Taq DNA polymerase, 40 mM Tris-HCl (pH 8.4), 100 mM KCl, 6 mM $MgCl_2$, 40 U ml^{-1} UDG and dNTP mix with dUTP (Invitrogen, Karlsruhe, Germany), 1 µl of ROX Reference Dye, 70 pmol of each *subIa/II* and 50 pmol of *nprI/IIa*, 5 µl of DNA, and MilliQ water to complete the 50 µl volume. The thermal cycler was programmed as: 2 min at 50°C for to activate uracil-N-glycosylase and 10 min at 95°C for to denature DNA and activate polymerase. This was followed by 40 cycles as followed: 20 s at 95°C for to denature DNA, 30 s at 53°C for for annealing of the primers and 30 s at 72°C for for extension.

2.3.7.3 TagMan 16S rRNA PCR assay

16S rRNA assay was performed in 50 µl as followed: 10 pmol of each primer (FP/RP 16S rRNA), 7.5 pmol of TaqMan probe (see Table 6), 10 nmol of dNTP mix with dUTP, 5 µl of 10×reaction buffer, 4.5 mM of $MgCl_2$ and 1.25 U of Ampli *Taq* Gold DNA Polymerase (Applied Biosystems, Foster City, USA), 0.5 µl of uracil-N-glycosylase, 5 µl of template DNA and MiliQ water to complete the 50 µl volume. The thermal cycler was programmed as

follows: 2 min at 50°C to activate uracil-N-glycosylase, 10 min at 95 °C to denature DNA and activate polymerase and finally 40 cycles of: 20 s at 95° C for to denature DNA and 60 s at 62 °C for annealing.

Table 6: 16S rRNA specific TaqMan- probe

Probe[1]	Composition (5´- 3´)[2]	Position (nt)[3]
probe 16S rRNA	TKCGCGTTGCDTCGAATTAAWCCAC	951- 975

[1] The probe is marked: on 5´- end FAM and on 3´- end TAMRA

[2] Base symbol: K = T/G, D = A/G/T, W = A/T

[3] Position in the 16S rRNA gene of *E. coli* (Brosius *et al.*, 1981)

2.3.7.4 Examination of sensitivity of Real-Time PCR and DNA extraction recovery

The sensitivity of the *npr/sub* assay was determined using series dilution (1:10, 1:100, 1:250, 1:500, 1:1000) of extracted soil DNA. To eliminate inhibition of PCR and no further to rise detection limit soil DNA extracts were diluted 1:250. The absence of inhibitory substances at the dilution used was confirmed by addition of 5 x 10^5 copies of the standard DNA in series of 10-fold diluted soil extracts. Specificity of SybrGreen assay was checked on agarose gel where 10% of Real- Time PCR products from one of three independent replicates of each soil samples were put on the gel (see section 2.3.5). The number of 16S rRNA, *sub* and *npr* gene copy was calculated per 1 g dry matter of soil.

To determine DNA extraction efficiency a spiked soil samples was analyzed by Real-time PCR. The bulk soil was amended with environmental DNA extracted from unamended soil samples. The *npr* copies number of amended DNA were ten times higher as the number of target gene in the unamended soils. The DNA was extracted in three parallels from the amended soil samples collected in October (see section 2.3.2.2). The number of respective gene (*npr*) was measured by Real-time PCR from non-inhibited dilutions (1:250).

2.3.8 Terminal restriction fragment length polymorphism (T-RFLP)

Terminal restriction fragment length polymorphism (T-RFLP) is a culture-independent, PCR based technique that allows the assessment of a diversity of complex bacterial communities using a DNA sequencer. T-RF patterns are obtained and analyzed in a series of steps that include PCR, restriction enzyme digestion and capillary gel electrophoresis. DNA extracted from a sample is subjected to PCR using a primer that is fluorescently labeled at one end. Therefore in PCR reaction resulting amplicons maintain terminal label on the one site of the molecule and are of a similar size. The amplified gene fragments are then digested with a selected restriction enzyme to separate amplicons that origin from different organisms and to obtain fingerprint. The gene sequences originated from various organisms containing different restriction sites resulting in terminally labeled fragments of different sizes. These terminal restriction fragments (T-RFs) are subjected to capillary gel electrophoresis, usually a DNA sequencer with a fluorescence detector so that only the fluorescently labeled terminal restriction fragments (T-RFs) are visualized. Principle of T-RFLP approach is shown in Figure 4.

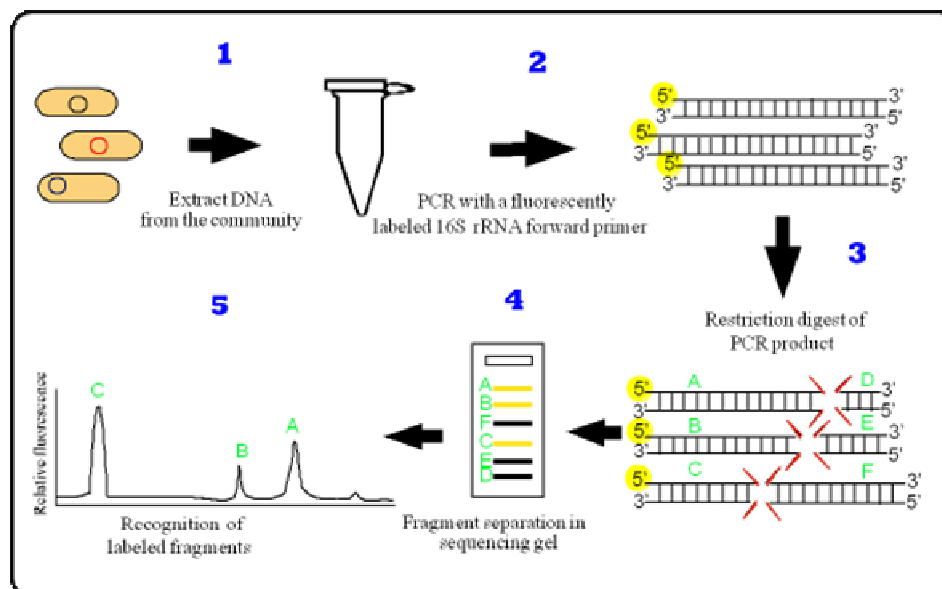


Figure 4: Principle of T-RFLP approach (http://rdp8.cme.msu.edu/html/t-rflp/02.html)

2.3.8.1 Selection of restriction enzymes

The selection of restriction enzymes that yields the highest number and most even distribution of T-RFs were of crucial importance to establish reliable method for comparing microbial communities and assessing community dynamic. NCBI database (http://www.ncbi.nlm.nih.gov/entrez) was used to retrieve known *sub* or *npr* sequences (see Tables 7 and 8). The sequence between two primer-binding sites in each retrieved genes was applied to NEBcutter V2.0 software (http://tools.neb.com/NEBcutter2/index.php) that accepts an input DNA sequence and produces a comprehensive report of the restriction enzymes that cleave the sequence. After comparing the variety of outputs of theoretical digests the most appropriate restriction enzymes to analyze *sub* and *npr* coding communities were found. HpyCH$_4$V (5`-TG▼CA-3´) found to be most appropriate for T-RFLP analysis of gene encoding for subtilisin whereas simultaneously use of HpyCH$_4$V, Alu I (5`-AG▼CT-3´), Sac II (5`-CCGC▼GG-3´) provided the best result for assessing *npr* community dynamic. To confirm the results obtained from the database analysis and to optimize T-RFLP conditions for reliable comparison of proteolytic bacterial communities in soil, three bacterial strains containing gene encoding for subtilisin and five bacterial strains possessing gene coding for neutral metalloprotease were tested.

Table 7: Origin of *sub* specific sequences utilized for NEBcutter analysis and expected fragment size after restriction endonuclease digestion of 319 bp long *sub*- amplicon

Bacteria	**NCBI accession number**	Cutting position (5´-3´) (bp)	Fragment size (bp)
		HpyCH$_4$V	Cy5-*sub Ia*
B. licheniformis	AF282893	70, 76, 100, 289, 304	70
B. licheniformis	S78160	70, 76, 289, 304	70
B. amyloliquefaciens	X00165	76, 172	76
Bacillus sp.	D29736	51, 79, 160	51
Geobacillis stearothermophilus	M64743	151	151
Bacillus sp.	U39230	205	205
B. subtilis subsp. natto	S51909	151, 172	151
Brevibacillus leterosporus	AY720895	43	43

NCBI = National Center for Biotechnology Information

Table 8: Origin of *npr* specific sequences utilized for NEBcutter analysis and expected fragment size after restriction endonuclease digestion of 233 bp long *npr*- amplicon

Bacteria	**NCBI accession number**	Cutting position (5´-3´) (bp)	Fragment size (bp)
		HpyCH$_4$V, Alu I, Sac II	Cy5-*npr IIa*
B. megaterium	X75070	82, 19, 76	151
B. thuringiensis	L77763	76, 127	106
B. cereus	M83910	19, 46, 76, 127	106
B. thermoproteolyticus	X76986	82, 85	148
B. stearothermophilus	M11446	51, 65, 114	119
Listeria monocytogenes	X54619	119	114
Alicyclobacillus acidocaldarius	U07824	51, 65	168
Lactobacillus sp.	D29673	19, 46, 76	157

NCBI = National Center for Biotechnology Information

2.3.8.2 PCR amplification and purification of the PCR products

Sub and *nprA* specific PCR amplification was done using the primers (*subIa/II* and *nprI/IIa*) and conditions describe in section 2.3.6 (see Tables 3. and 5.) where *sub Ia* (forward primer) and *npr IIa* (reverse primer) were fluorescently labeled. Cy5 fluorescence dye (agitation maximum ca. 600 nm) was used to mark off the primers. MinElute PCR Purification Kit (Qiagen, Hilden, Germany) was used to purify PCR products (see section 2.3.6.2). Quantification of the amplicons was done by PicoGreen assay (see section 2.3.3).

2.3.8.3 Restriction endonuclease digestion

Purified PCR products were digested using New England Biolabs restriction endonuclease (New England Biolabs, Massachusetts, USA). Digestion of *sub* specific PCR products was set in 20 µl volume using the following protocol.

NEBuffer 4	3 µl
HpyCH$_4$V (5 U µl^{-1})	0.2 µl
PCR products	120 ng
dH$_2$O adjust vol. to	20 µl

Digestion was incubated at 37°C for 14 h.

Digestion of *nprA* PCR products was done in 20 µl volume as followed.

NEBuffer 4	3 µl
$HpyCH_4V$ (5 U $µl^{-1}$)	0.4 µl
Alu I (10 U $µl^{-1}$)	0.2 µl
Sac II (20 U $µl^{-1}$)	0.1 µl
PCR products	120 ng
dH_2O adjust vol. to	20 µl

Digestion was incubated at 37°C for 3 h.

The digested *sub* and *npr* PCR products were immediately purified using Qiaquick Nucleotide Removal Kit (see section 2.3.6.3). Quantification of the digested products was done by PicoGreen assay (see section 2.3.3).

2.3.8.4 Detection and analysis of *sub* and *npr* T–RFs

Detection and analysis of fluorescently labeled T-RFs were realized on automated sequence analyzer, CEQ™ 2000 XL DNA Analysis System, (Beckman Coulter GmbH, Fuellerton, USA). *Sub* or *nprA* digested products were mixed with CEQ SLS buffer and 600 bp DNA standard (Beckman Coulter GmbH, Fuellerton, USA) that contains fragments of defined size. Reaction approach was done in 30 µl as followed.

CEQ SLS Buffer	26.6 µl
CEQ DNA Size Standard Kit – 600	0.4 µl
DNA (10 ng)	3 µl

Prior to running on capillary electrophoresis system one drop of oil was added to each of the reaction.

Detection of T-RFs was set using the conditions given in Table 9.

Table 9: Program cycle conditions of T-RFLP analysis on CEQ™ 2000 XL DNA Analysis System

Step	Reaction	Setting parameters
1	Capillary heating	50°C
2	DNA denaturation	90°C 120 s
3	Injection of probes to capillary	2.0 kV 20 s
4	Separation of probes in capillary	3.5 kV 180 min

2.3.8.5 Analysis of T-RFLP data

The T-RFLP raw data were automatically converted to a digitized form using CEQ™ 8000 Genetic Analysis System (Beckman Coulter GmbH, Fuellerton, USA). Although the program made a peak table with all peaks in electropherogram, each electropherogram was checked manually to ensure that only true peaks are in the table (and only once) and to omit detection of the background. Peaks heights with less than 1% of the summarized height of all peaks present in the electropherogram were considered as background and excluded from further analysis as well as T-RFs that were smaller than 50 bp. Additionally, uncut fragments were skipped from the analysis. They comprised on average 5.9% of the summed peak height. To standardize the results the height of each peak was divided with the summarized height of all peaks in one electropherogram to obtain the relative abundance of respective peak.

2.3.9 Cloning analysis

2.3.9.1 Design of reverse *npr* PCR primer for phylogenetic analysis

In order to reveal much phylogenetic information about the *npr* bacterial proteolytic community, new reverse primer was constructed to amplify longer *npr* fragment. Forward primer used were describe elsewhere (Bach *et al.*, 2001). For primer design, nucleotide sequences of genes of 61 organisms coding for a neutral metalloprotease were aligned using ClustalW (Thompson *et al.*, 1994). The alignment was manually screened for a conserved region, which could be targeted by the new reverse primer. The expected size of the amplified fragment is 483 bp.

2.3.9.2 Cloning

Purified *nprB* PCR products (see section 2.3.6.1) were cloned into pCR®2.1 vector of TA Cloning Kit (Invitrogen, Karlsruhe, Germany) as described below.

Purified *nprB* PCR product	3 µl (approx. 150 ng)
10X ligation buffer	1 µl
Sterile water	3 µl
pCR®2.1 vector (25ng/µl)	2 µl
T4 DNA ligase	1 µl

Ligations were incubated at 14°C overnight. 2 µl of ligation mixture was transformed into chemically competent One Shot® TOP10 cells (Invitrogen, Karlsruhe, Germany) provided in the kit following manufacturer's instructions. Colonies were inoculated in LB medium (see section 2.3.1.1) containing 50 µg ml^{-1} of kanamycin and 40 µl of 40 mg/ml stock X gal solution (Qbiogene, Germany). Plasmids were isolated using NucleoSpin Plasmid Mini Kit (see section 2.3.2.3). Purified plasmids were tested for inserts by *Eco*RI digestion (MBI Fermentas, Heidelberg, Germany). Digestion was set using the following protocol.

10x Buffer O^+	2 µl
Plasmid	4 µl
Restriction enzyme *Eco*RI (10 U μl^{-1})	0.5 µl

Nuclease-free water	13.5 µl

Digestion was incubated at 37°C for 2 h. The digested products were immediately checked on 1% agarose (see section 2.3.5) at 100 V for 1 h.

2.3.9.3 Sequencing

Inserts were sequenced on ABI PRISM® 3730 DNA Sequencer (Applied Biosystems, Foster City, USA) using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, USA). Reactions were performed as described below.

BDT Buffer	1 µl
M13 reverse Sequencing Primer (10 pmol)	1 µl
Template	100 ng
dH_2O adjust vol. to	4 µl
BigDye Terminator v3.1 Mix	1 µl

Thermal cycling program followed was: 96°C for 1 min, 48°C for 20 sec, 60°C for 1 min for 50 cycles.

2.3.9.4 Purification of sequencing reaction

Removal of unincorporated dye terminators directly from sequencing reaction was done using DyeEx 2.0 Spin Kit (Qiagen, Hilden, Germany). Prior to purification procedure sequencing reactions was diluted in 15 µl of LiChrosolv water (Merck and Co.,Inc, Whitehouse Station; USA). To cleanup sequencing reactions manufacturer's instructions was followed. Before loading on the sequencer 25 µl of LiChrosolv water was added to the pure products.

2.3.9.5 Sequence analysis

Obtained nucleic acid sequences were translated to amino acids using ExPASy translate tool (http://www.expasy.org/tools/dna.html) and compare to sequences deposit in GenBank databases using NCBI BLAST (http:www.ncbi.nlm.nih.gov/BLAST/). Sequences that were not characterized as proteases were excluded from further analysis. Alignments were made

using Workbench 3.2 software (http://workbench.sdsc.edu/). Phylogenetic trees were constructed using program in the Phylip software package version 3.65. Calculation of distance matrices of the amino acid sequences was carried out by the PROTDIST programs. Trees were inferred using NEIGHBOR program, which implements the neighbor-joining method of Saitou and Nei (1987) and the UPGMA method of clustering. Phylogenetic trees from amino acid sequences were also constructed with PROPARS (parsimony) and PROML (maximum likelihood) to check for the stability of tree topology. Bootstrap analysis with the SEQBOOT program was used to estimate the reliability of phylogenetic reconstructions (1000 replicates), constructing trees with PROTDIST and NEIGHBOR. The consensus tree was generating with the CONSENSE program. Treeview 32 software was applied to view the trees.

2.3.9.6 Nucleotide sequence accession numbers

The clone sequences determined in this study have been submitted to GenBank under accession numbers EF152076-EF152168.

2.3.9.7 Assignment of cloned *npr* sequences to T- RFLP

To connect the cloned *npr* sequences to *npr* T-RFs, computer based simulation was made. Obtained cloned *npr* sequences were cut to get 233 bp long fragments. They were subjected to NEBcutter V2.0 software (http://tools.neb.com/NEBcutter2/index.php) to search for the recognition sites of HpyCH$_4$V, AluI and SacII that were used for T-RFLP assay. All sequences that were >98% similar were analyzed only once.

2.4 Microbial activity measurements

2.4.1 Proteolytic activity analysis

Activity of soil proteases (EC 3.4.21-24) was measured by the modified procedure of Ladd and Butler (1972). The assay is based on the cleavage of casein, used as a substrate analogue, into trichloroacetic acid-soluble peptides containing tyrosine and tryptophan residues. These peptides react with chromogenic compounds that can be spectrophotometrically detected and measured.

Soil samples (1 g) were incubated for 2 h at 50°C with addition of 5 ml of 2% sodium caseinate (Merck, Darmstadt, Germany). At the end of incubation 5 ml of 2% sodium caseinate was added to the blank probes (control) followed by 5 ml of 0.92 M trichloroacetic acid. The aromatic amino acids released in soil suspensions were filtered and placed in the test tube containing 7.5 ml of alkaline reagent mixture (50 ml sodium carbonate, 1 ml of 20 mM copper sulfate pantahydrate, 1 ml of 35 mM potassium sodium tartrate tetrahydrate). Finally, 5 ml of Folin-Ciocalteu reagent (Sigma-Aldrich Chemie, Steinheim, Germany) was added and after 1.5 h of incubation at room temperature absorption was measured at 700 nm by spectrophotometer (CE 3021, Cecil Instruments Ltd., Cambridge, England). From the subtracted average value from tyrosine standards the standard curve was generated allowing estimation of potential proteolytic activity of the soil samples. Proteolytic activity was reported as µg tyrosine g dm^{-1} $(2h)^{-1}$.

2.5 Statistical analysis

Variables were tested for normality. Data that were not normally distributed were log- or quadrate square transformed to obtain Gaussian distribution. A Pearson correlation analysis was used to check for correlation between 16S rRNA, *sub* and *npr* copies number, proteolytic activity, organic C, total N, WEON, WEOC and water content. T-RFLP profiles were compared using multivariate statistical methods. Principal component analysis (PCA) is performed on the symmetric correlation matrix using R-project software (Team, 2005) with the aim of identifying the samples that show similar patterns. Prior to multivariate analysis, variables were median centered. Effects of time, site and depth on components in PCA were

tested using ANOVA analysis supplemented in R-Project software. All analyses were done at $P<0.05$.

3. Results

3.1 Abiotic soil parameters

From the field A15 the soil samples from four different soil types Gleyic Cambisol (GC); Luvisol/ Cambisol (LC); Cambisol/ Pseudogley (CP) and Gley/ Kolluvisol (GK) and three soil depths: D1 (≈0-10 cm), D2 (≈10-30 cm) and D3 (≈30-40 cm) were sampled.

The pH was slightly acidic ranging from 4.4 to 5.8. pH variations between the investigated sites were very low. Depth effect was visible only at GC and CP at D3 where the lowest pH was measured whereas at LC and GK no depth effect was noticed (Table 10).

Measured amount of C_{org} was between 0.2 and 1.6%. The contents of C_{org} decreased with soil depth. Comparing the C_{org} content at different sites it was found that highest C_{org} content was measured at GK, whereas the lowest values was noticed at LC (Table 10).

The concentration N_{tot} ranged from 0.02 to 0.16%. The maximum of N_{tot} was observed at D1 and minimum at D3. A site specific effect was also visible: the highest amount of N_{tot} was measured at GK and the lowest at LC (Table 10).

The amount of particles smaller than 2 mm varied at different sites or depths. The highest proportion of particles was noticed at GK whereas at GC the lowest values were measured. The amount of particles smaller than 2 mm was not consistent at different depths: at GC the lowest proportion was found at D3, whereas at LC, CP and GK the minimum values were noticed at D2 (Table 10).

The clay content was not consistent at different depths: at GC and LC it decreased with increasing soil depth, whereas at CP and GK the clay content was highest in deepest soils. At CP the maximum and at LC the minimum of clay particles was measured (Table 10).

The silt content differed with different depths and sites. At GC and LC the lowest amount of silt was noticed at D3 whereas at CP and GK no depth effect was visible. Comparing the silt content at different sites, the reduced silt amount at LC and maxima at GK was found during the whole period of investigation (Table 10).

At GC and LC the sand content was highest at D3 whereas at CP and GK the lowest amount of sand was observed in deepest soils. Beside soil depth, site-specific effects were also visible. The sand content was lowest at GK whereas the highest amount of sand was measured at LC (Table 10).

Table 10: Some physicochemical properties of the soil horizons sampled at the investigated field

soil type	horizons *	depth cm	depth number	pH ($CaCl_2$)	Clay [%]	Silt [%]	Sand [%]	Particles>2mm [%]	organic C %	total N %
GC[a]	Ap	0-10	D1	5.8	22	36	42	2	1.4	0.15
	Awp	10-40	D2	5.7	22	40	38	1.7	0.9	0.10
	Awp	30- 40	D3	4.7	18	22	60	0	0.3	0.04
LC[b]	Ap1	0-10	D1	5.5	14	25	61	1.6	1.1	0.11
	Ap2	10-23	D2	5.6	14	24	62	0.3	0.7	0.08
	IIB(b)tv	23-40	D3	5.7	5	5	90	0.9	0.2	0.02
CP[c]	Ap	0-10	D1	5.7	19	49	32	4.6	1.4	0.14
	SwAp	10-27	D2	5.6	20	47	33	1.5	1.0	0.10
	BvSw	27-40	D3	4.4	27	47	26	2.4	0.3	0.04
GK[d]	Ap1	0-15	D1	5.5	16	53	31	3.3	1.6	0.16
	Ap2-M	15-30	D2	5.8	18	52	30	3.1	1.1	0.11
	Go-M	30-40	D3	5.7	22	56	22	4.6	0.6	0.07

* according to the FAO soil classification system (Fao, 1976)

[a-d] soil types: [a] Gleyic Cambisol, [b] Luvisol/ Cambisol, [c] Cambisol/ Pseudogley, [d] Gley/ Kolluvisol

The concentration of WEOC as a measure of carbon bioavailability, decreased by soil depth, with a maximum observed at D1 and minimum at D3. Comparing the WEOC content at different sites it was found that in April the highest WEOC content was measured at GC, in July at CP and in October the maximum WEOC was noticed at GK. Seasonal effects differed for the different sites: at GC and GK maximal WEOC were measured in April and at LC and CP in July whereas overall decrease in WEOC amount was measured in October (Figure 5).

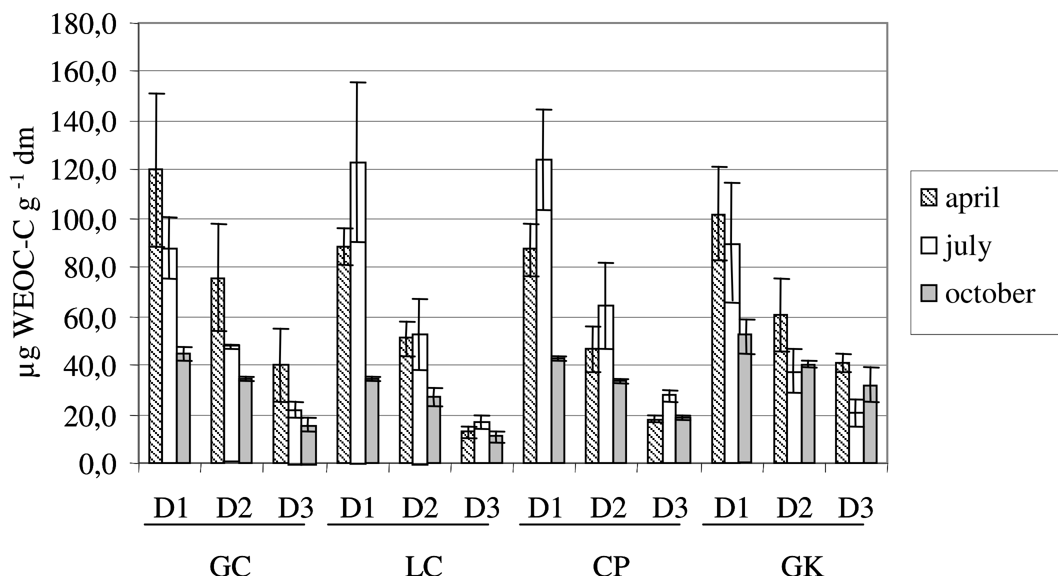


Figure 5: WEOC measurements investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard deviation of means.

For the period of April no data about WEON as a measure of bioavailable portion of nitrogen, was available. The measurements of WEON in July and October revealed the highest WEON content at D1 and the lowest at D3. The highest value was noticed in July at LC and in October at GK. Temporal changes in WEON amount were not consistent for sites: at GC and LC WEON content reached maxima in July whereas at CP and GK highest values were measured in October (Figure 6).

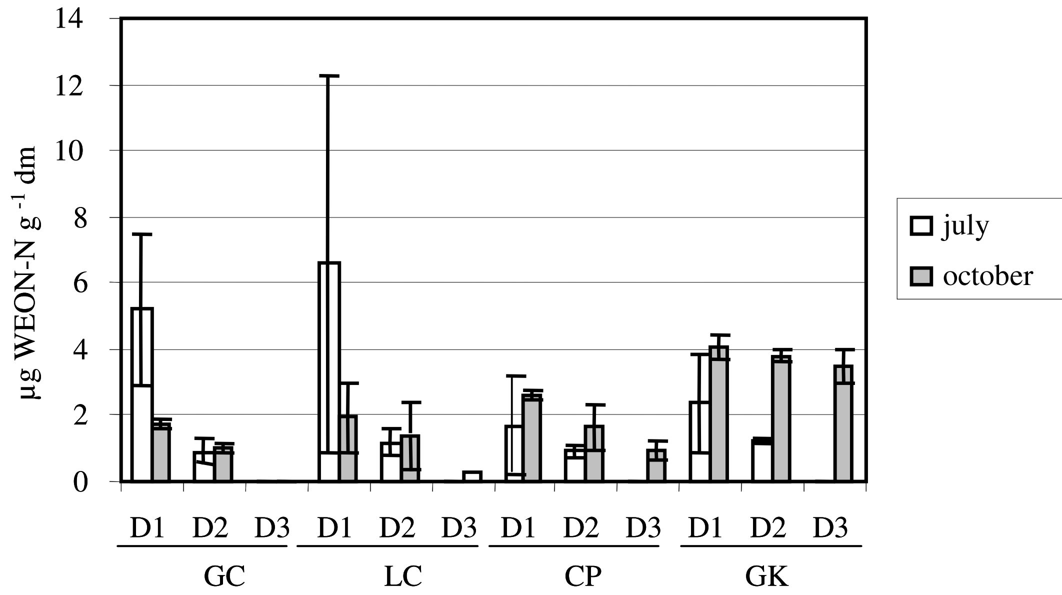


Figure 6: WEON measurements investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard deviation of means.

3.2 Isolation of DNA from soil samples

The amount of extracted DNA was between 24 and 211 ng/µl reaching the highest values in upper soils (Figure 7). Beside soil depth, site-specific effects were also visible. The highest concentration of DNA was measured at GK during the whole period of investigation. Temporal fluctuation was not consistent at different sites: at GC maxima of DNA concentration was noticed in October, at LC and CP in July whereas at GK no differences between July and October were observed.

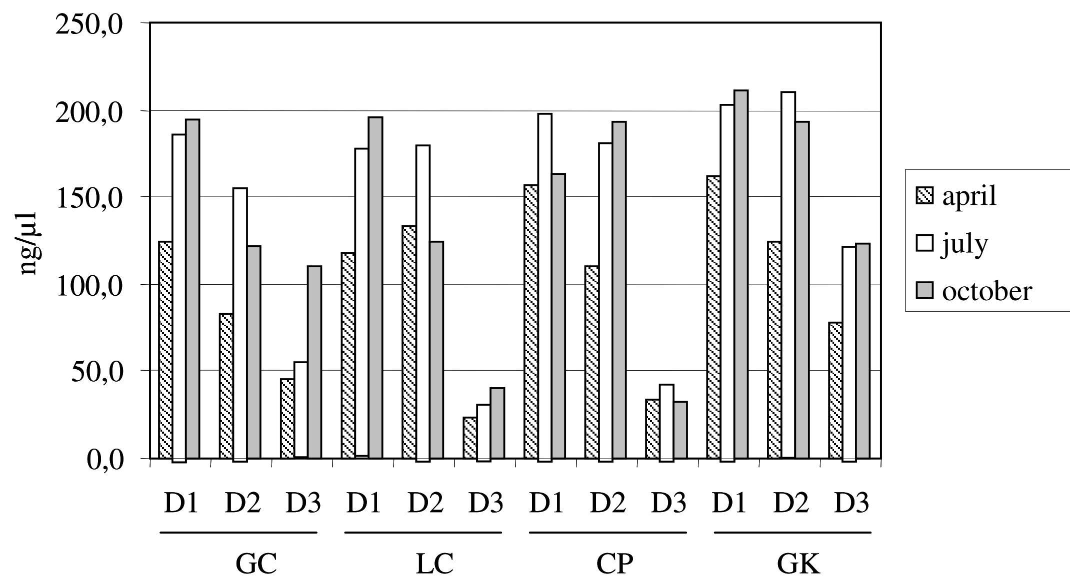


Figure 7: DNA concentration measurements investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003.

3.3 Products of soil *npr* and *sub* PCR amplification

The amplification of the part of the neutral metalloprotease´s and subtilisin´s genes using specific primers, *nprI/IIa* and *subIa/II*, brings the PCR products as shown in Figure 8 and 9 for October samples. The primer binding site and specificity of the used primers was described by Bach *at al.* (2001).

Amplification of the part of the gene encoding for neutral metalloprotease resulted in the amplicons that length was 233 bp (Figure 8). *Npr* specific PCR products were amplified from each of the four investigated soil types and three soil depths in April, July and October.

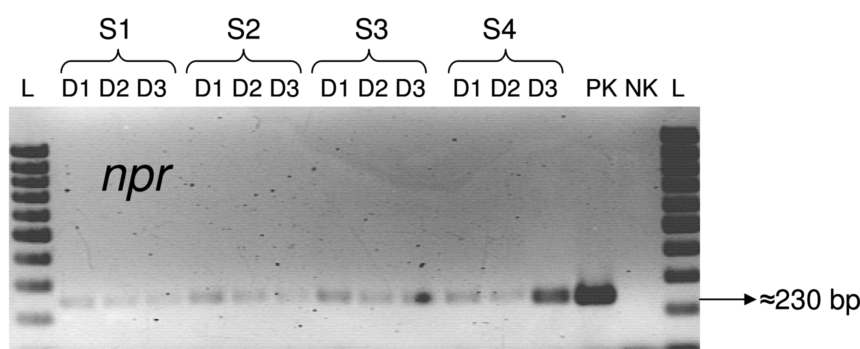


Figure 8: Obtained *nprA* PCR products in soil samples with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) electrophoretically separated on 1.5 % agarose gel. As an exemplar results from October were shown.

Outcome of the amplification of the part of the gene encoding for subtilisin were 319 bp long PCR products (Figure 9). *Sub* specific amplicons were amplified from each of the four investigated soil types and three soil depths in April, July and October.

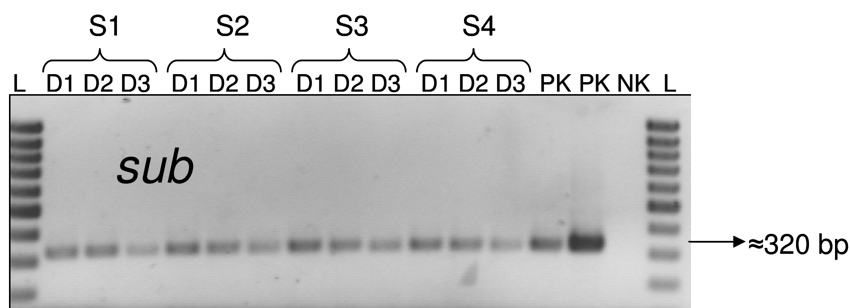


Figure 9: Obtained *sub* PCR products in soil samples with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) electrophoretically separated on 1.5 % agarose gel. As an exemplar results from October were shown.

3.4 Quantification of *npr*, *sub* and 16S rRNA

3.4.1 Plasmid standard preparation and calculation of gene copies number

Standard curve was generated from a dilution series constructed from a standard DNA. To obtain standard DNA, PCR was applied to amplify the part of the *sub, nprA* and 16S rRNA gene using *subIa/II, nprI/IIa* and 16S rRNA primers. For that purpose genomic DNA from *Bacillus subtilis*, *Bacillus cereus* and *E. coli* was utilized. The PCR products were visualized on the agarose gel, purified and cloned. Ten transformants were picked and subjected to high yield plasmid isolation. After *Eco*RI endonuclease digestion plasmids that contain 319 (*sub*), 233 (*npr*) or 263 (16S rRNA) bp long inserts (+20 bp vector sequence) were sequenced to check for the specificity. The received sequences were compared with known protease or 16S rRNA sequences from the public databases to confirm the origin. Quantification of the plasmid DNA was done by PicoGreen assay. Concentrations of the standards DNA were as followed: 220 ng/µl (*sub*), 215 ng/µl (*npr*) and 206 ng/µl (16S rRNA). To estimate the copies number of unknown probes in the assay it was necessary to calculate the number of the gene copies for standard DNA. The calculation was done as followed:

3851 bp (vector length) + 319 bp (*sub* insert) = 4170 bp

4170 bp x 660 g/mol (molecular weight of one base pair) = 2.75×10^{6} g/mol

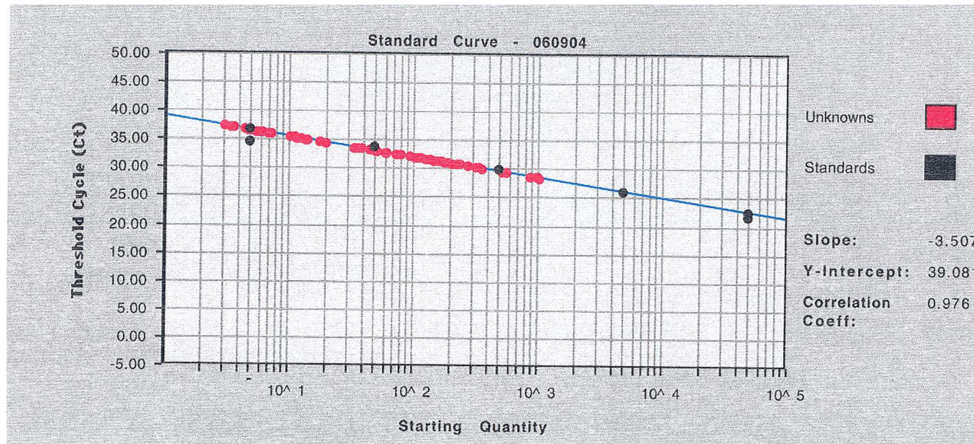
220×10^{6} g/ml (concentration of the *sub* standard DNA) / 2.75×10^{6} g/mol = 8×10^{-11} mol/ ml

In one mol there are 6.022×10^{23} molecules that mean 4.82×10^{13} molecules in 8×10^{-11}-mol/ml of *sub* standard DNA. Therefore in 1 µl of *sub* standard DNA there were 4.82×10^{10} molecules and equal number of *sub* specific inserts. The same calculation was followed to estimate the *npr* and 16S rRNA standard copies number with the respect to different lengths of the respective inserts.

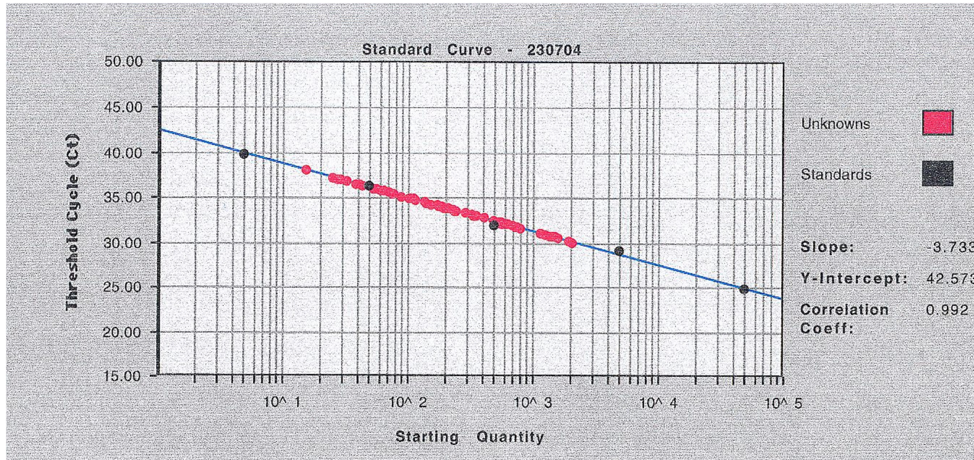
3.4.2 Examination of Real-Time PCR detection limits and DNA extraction efficiency

Tenfold serial dilution of the plasmids ranging from 10^{1} to 10^{5} was used as template, by duplicate, to determine the calibration curves. The standard curves had a linear range over

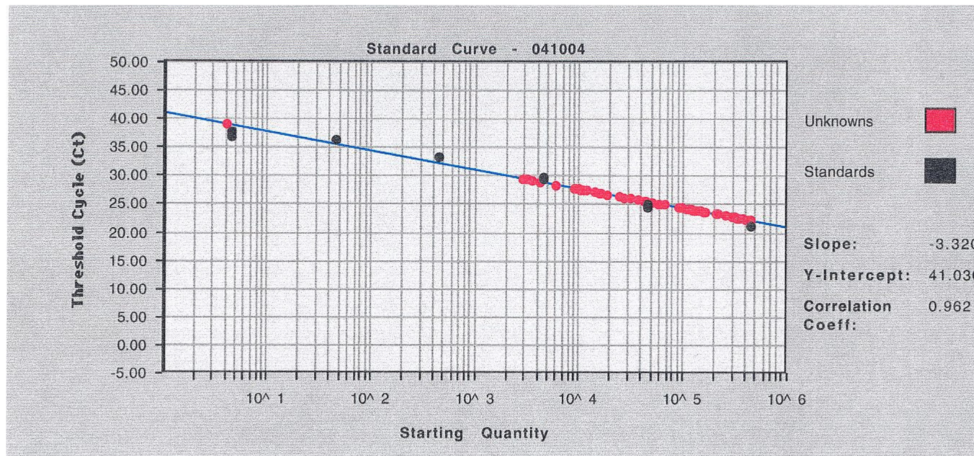
five orders of magnitude for each of the targets with correlation coefficients above 0.950 (Figure 10).



a)



b)



c)

Figure 10: Standard curves for quantification of soil a) *npr*, b) *sub* and c) 16S rRNA. Black dots represent standard DNA with known copies number and rot represent soil DNA with unknown gene copies number.

Environmental DNA was serially diluted to see for inhibitory effects of substances coextracted during DNA extraction to Real-Time PCR assay. No amplification was detected when non-diluted DNA or 1:10 dilution was applied. An enhancement of PCR effectiveness could be observed when DNA eluates were diluted up to 1:250. To eliminate inhibition and not to further raise the detection limit, 1:250 dilutions of the soil DNA were used as template. In addition, 1×10^5 copies of the standard DNA were added in serial dilutions of environmental DNA to check for Real-Time PCR inhibition. No inhibitory effect was observed in the dilution 1:250, which was used in present investigation. The detection limit of the assay was 2×10^5 targets per gram dry matter.

To determine DNA extraction efficiency a spiked soil samples was analyzed by Real-time PCR under the same conditions as unspiked soil samples. Real-Time PCR quantification indicated that target DNA recovery from investigated soil samples ranged from 36 to 62% (Table 12).

Table 11: DNA extraction efficiency of soil samples

	Recovery %											
Soil samples	GC D1	GC D2	GC D3	LC D1	LC D2	LC D3	CP D1	CP D2	CP D3	GK D1	GK D2	GK D3
mean ±S.D.	56 ±18	42 ±11	62 ±10	49 ±22	36 ±10	52 ±0.2	44 ±8.3	60 ±3.9	53 ±9.9	50 ±3.1	45 ±4.6	53 ±22

S.D. as percent of mean

3.4.3 Abundance of *npr*, *sub* and 16S rRNA gene copies in soil samples

The abundance of *npr, sub* and 16S rRNA copies determined at four different sites and three depths in April, July and October are shown in Figure 11. The abundance of 16S rRNA copies decreased with increasing soil depth during the whole period of investigation. The values ranged between 1 x 10^8 (LC D3 in April) and 2 x 10^{10} (CP D1 in July). Beside soil depth, site-specific effects were also visible. Abundance of 16S rRNA copies was constantly lowest at LC whereas maximum number of 16S rRNA at different sites varied in time. The highest number of 16S rRNA was measured in April at GC, in July at CP and in October at GK. Temporal fluctuation was not consistent at different sites: at GC maxima of 16S rRNA was noticed in April, at CP in July whereas at LC and GK in October.

Npr copies number showed a constant decrease with increasing soil depth, with D3 having the lowest and D1 the highest number of *npr*. *Npr* copy number reached the maximum at GK D1 in July (1 x 10^7) and minimum at LC D3 in April (less than 2 x 10^5 targets per gramm dry matter). Comparing the *npr* number at different sites, the clear tendency in reduced *npr* numbers at LC and maxima at GK during the whole period of investigation was found, following the same tendency in the nutrient and clay/silt content (Table 11). Seasonal effects differed for the different sites: at GC, CP and GK maximal *npr* numbers were measured in April whereas at LC *npr* copy number reached the maximum in July.

During the whole period the highest abundance of *sub* was measured at D1 and the lowest at D3. The highest value was noticed at GK D1 in October (3 x 10^7) and lowest at LC D3 in April (less than 2 x 10^5 targets per gramm dry matter). Observation of *sub* number at different sites showed clear reduction of abundance at LC during the whole period, supporting the findings of lower 16S rRNA and *npr* number at LC. Maximal number of *sub* was measured at GK (except April where the highest *sub* number was at GC). Temporal changes in *sub* number were not consistent for sites: at GC the *sub* number reached maxima in April, at LC and CP in July whereas at GK highest values were measured in October.

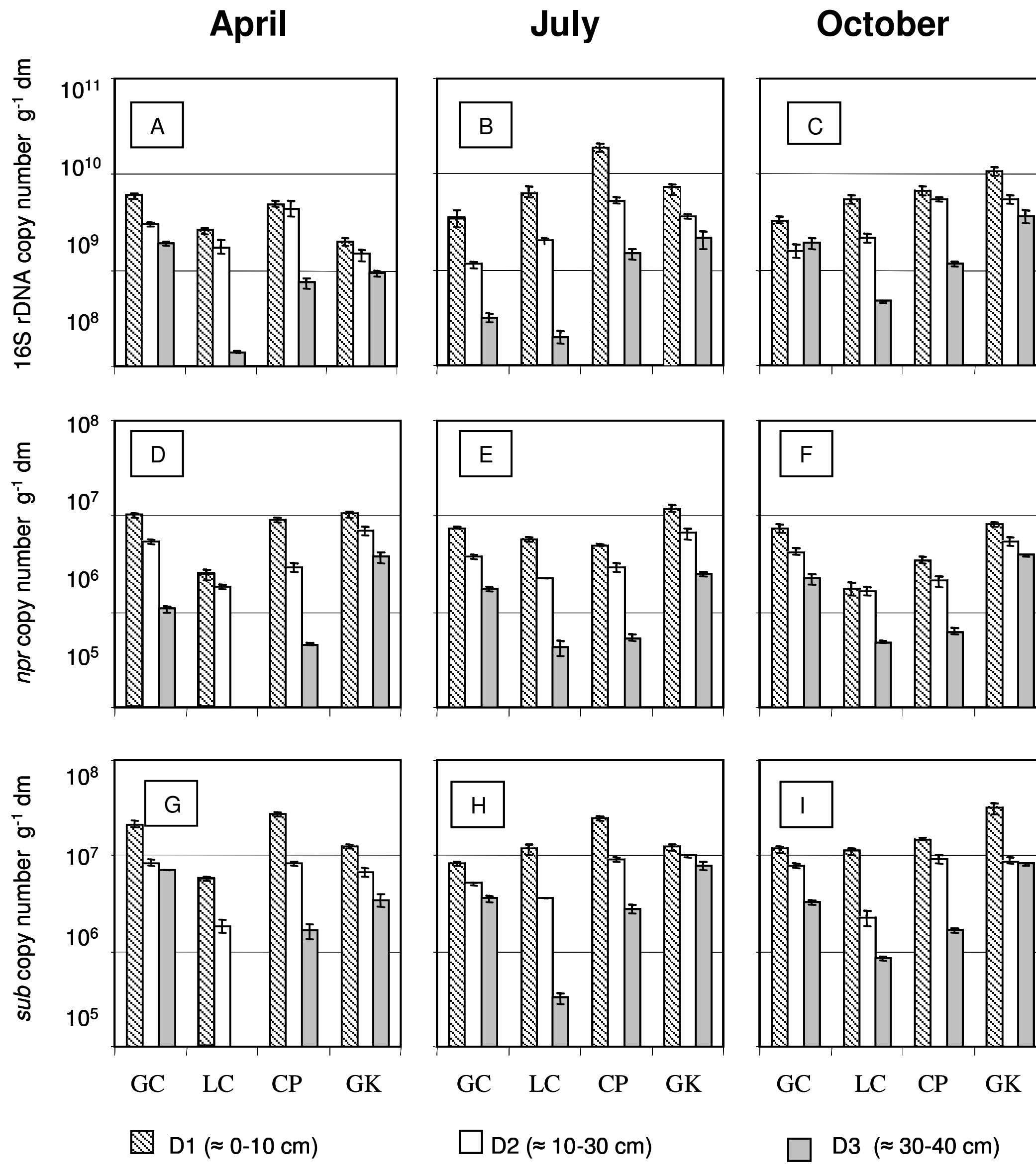


Figure 11: Number of total (A-C), *npr* (D-F) and *sub* (G-I) coding proteolytic bacteria investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard errors of means.

3.5 Proteolytic activity potential

Figure 12 shows the potential activity of proteases measured in the respective soil samples. Higher dynamics of proteolytic activity were observed in topsoil in comparison to subsoil and a constant decrease of activity with increasing depth was observed. Proteolytic activity was highest at GK D1 in October (272 µg tyrosine g dm^{-1} $(2h)^{-1}$) and lowest at LC D3 in April (not detectable). During the whole investigation period the highest activity was measured at GK and the lowest value was registered at LC. Potential proteolytic activity showed maxima at GC, LC and GK in October and at CP in April. In July there was an overall decrease in activity.

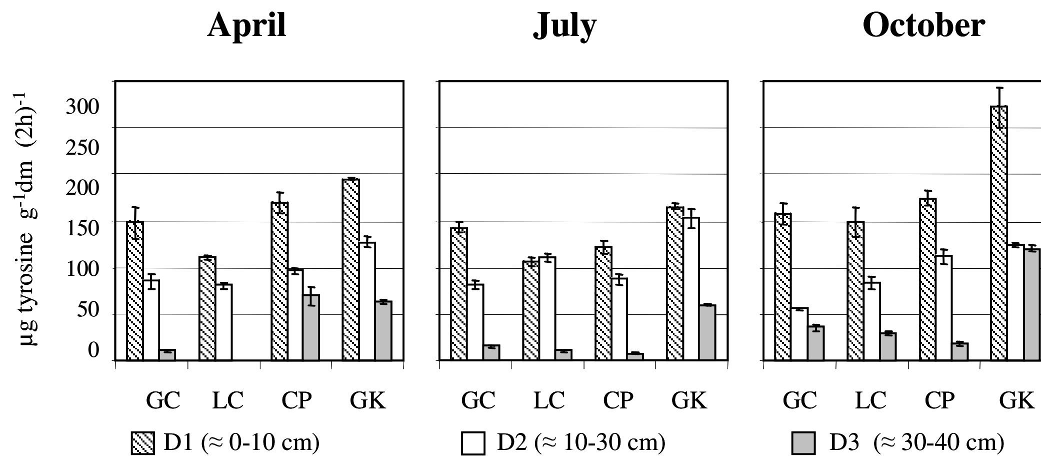


Figure 12: Potential proteolytic activity investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard errors of means.

3.6 Relation between soil bacteria protease genes, protease activity and 16S rRNA numbers

Positive and significant correlation (P<0.05) was calculated between *sub* copy number and 16S rRNA copy number, *npr* copy number, proteolytic activity, organic C, total N and WEOC (Table 13). Significant (P<0.05) but slightly correlation was observed between *sub* copy number and WEON and no correlation between *sub* number and water content was detected (Table 13).

Npr copy number was positively and significantly (P<0.05) correlated with *sub* copy number, proteolytic activity, organic C, total N and WEOC and no correlated with water. Significant but slightly correlation was observed in comparison to 16S rRNA copy number and WEON (Table 13).

Table 12: Pearson correlation matrix between microbiological and chemical properties in soil samples from field site with four different sites and three different depths in April, July and October 2003.

	16S	*sub*	*npr*	activity	C_{org}	Ntot	WEOC	WEON	water
16S	-	0.756**	0.356**	0.563**	0.602**	0.581	0.514**	0.331**	-0.192*
sub	0.756**	-	0.660**	0.750**	0.757**	0.749**	0.551**	0.410**	-0.073
npr	0.356**	0.660**	-	0.733**	0.822**	0.833**	0.665**	0.463**	-0.018
activity	0.563**	0.750**	0.733**	-	0.896**	0.888**	0.511**	0.521**	-0.197*
Corg	0.602**	0.757**	0.822**	0.896**	-	0.994**	0.697**	0.518**	-0.205*
Ntot	0.581	0.749**	0.833**	0.888**	0.994**	-	0.698**	0.514**	-0.163
WEOC	0.514**	0.551**	0.665**	0.511**	0.697**	0.698**	-	0.435**	-0.131
WEON	0.331**	0.410**	0.463**	0.521**	0.518**	0.514**	0.435**	-	-0.096
water	0.192	0.073	0.018	0.197*	0.205*	0.163	0.131	0.096	-

* P < 0.05, ** P < 0.01

Significant overall positive correlation between *sub* or *npr* gene number and potential proteolytic activity was found (Table 13). However when the relationship between both measured parameters in two entirely diverse soil samples was closely investigated significant positive relationship between *sub* or *npr* copy number and potential proteolytic activity was found only in LC D3 soil sample where highest sand content was measured (Figure 13). R^2 was 0.94 for *sub*/activity relationship and 0.74 for *npr*/activity, respectively. No relation occured between *sub* and activity at CP D3 where highest clay content was measured (R^2= 0.10). The relationship between *npr* and activity at CP D3 was very weak (R^2= 0.34).

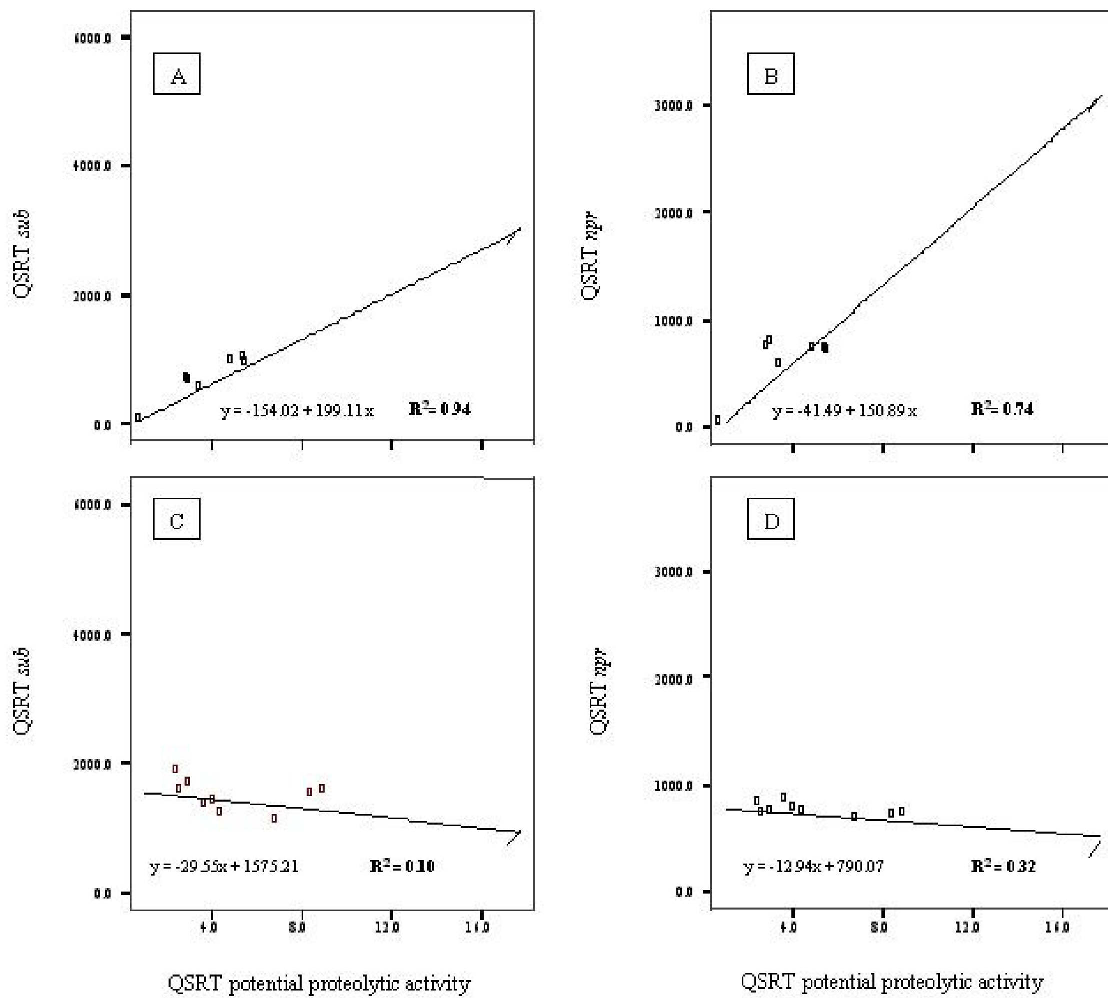


Figure 13: Relationship between *sub* copy number and potential proteolytic activity at site/depth LC D3 (A) and site/depth CP D3 (C). Relationship between *npr* copy number and potential proteolytic activity at site/depth LC D3 (B) and site/depth CP D3 (D) in April, July and October 2003. (SQRT=square root transformed data)

3.7 T-RFLP analysis

3.7.1 Optimization and specificity confirmation of T-RFLP analysis

The computer-simulated analysis suggested that sequence differences among T-RF lengths could be used to characterize complex proteolytic microbial communities. To confirm these findings and to establish the best conditions for T-RFLP analysis, three bacterial strains containing gene encoding for subtilisin and five bacterial strains possessing gene coding for neutral metalloprotease were tested. To completely digest PCR products and to obtain just one peak from one strain in the electropherogram, concentration of restriction endonuclease as well as duration of digestion were varied. The *sub* PCR products were digested with 1, 2 or 3 U of HpyCH$_4$V for 2, 3, 4 or 14 h. 1U of the respective enzyme and 14 h digestion allowed differentiation of three bacterial strains used as a model (Figure 14).

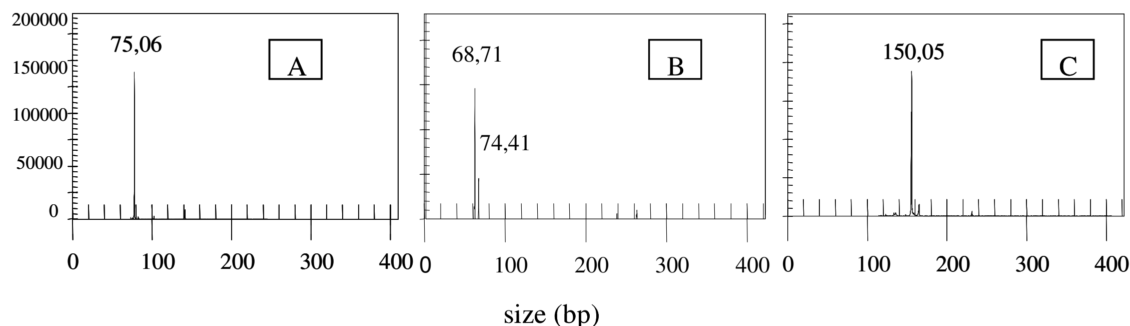


Figure 14: Electropherogram of the *sub* T-RFLP from (A) *Bacillus amyloliquefaciens* DSM 7, (B) *Bacillus licheniformis DSM* 13, (C) *Bacillus subtilis* DSM 10

The empirically obtained *sub* T-RFs were not completely matched with expected. The discrepancies were <1 bp (Table 14) and therefore neglectable because a difference of ± 2 bp in the sizes of T-RFs can occur due to nature of the gel separation.

T-RFs pattern of *B. licheniformis* consists of one double peak that is probably due to the quite nearness of two restriction sites (70 and 76) in *sub* amplicon that origin from *B. licheniformis* (see Table 7).

Table 13: Expected and observed lengths of *sub* T-RFs for the tree organisms possessing gene coding for Sub

Expected (observed) *sub*-T-RFs length (bp)		
Bacillus amyloliquefaciens	*Bacillus licheniformis*	*Bacillus subtilis*
76 (75)	70 (69, 74)	151 (150)

The *npr* PCR products were digested with either 1, 2 or 4 U of AluI, $HpyCH_4V$ and SacII for 1, 2, 3 or 14 h. 2U of the respective enzymes and 3 h digestion allowed differentiation of five bacterial strains used as a model and only one peak that originate from one bacteria strain was observed in the electropherogram (Figure 15).

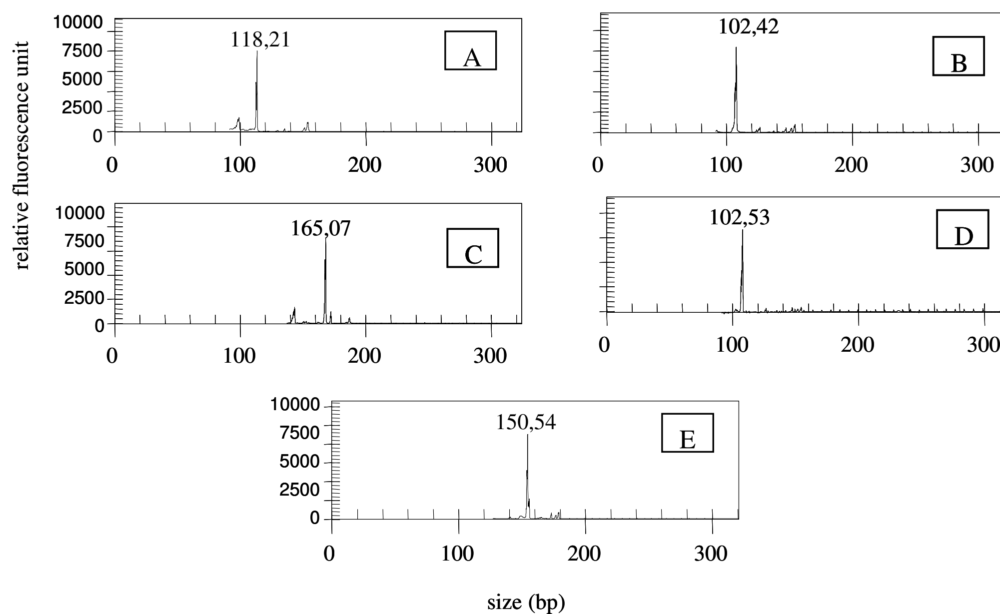


Figure 15: Electropherogram of the *npr* T-RFLP from (A) *Bacillus stearothermophilus* DSM 22, (B) *Bacillus thuringiensis* DSM 2046, (C) *Bacillus sp*. DSM 405, (D) *Bacillus cereus* DSM 3101, (E) *Bacillus megaterium* DSM 90

There were discrepancies between expected results based on the database analysis and empirically obtained *npr* T-RFs results. The differences were noticed for *B. cereus* and *B. thuringiensis* and were up to 4 bp (Table 15).

Table 14: Expected and observed lengths of *npr* T-RFs for the five organisms possessing gene coding for Npr

Expected (observed) *npr* T-RFs length (bp)				
B. stearothermophilus	*B. cereus*	*Bacillus sp.*	*B. thuringiensis*	*B. megaterium*
119 (119)	106 (102)	no data (165)	106 (102)	151 (151)

3.7.2 Evaluation of soil T-RFLP assays

3.7.2.1 *Npr* T-RFLP evaluation

Analysis of *npr* coding bacterial community investigated in soil samples from field site with four different soil types (GC,LC,CP, and GK) and three soil depths (D1-D3) in April, July and October provided 7-15 *npr* T-RFs that heights were above 1% of the heights of all peaks present in the electropherogram. *Npr* T-RF lengths ranged from 70 to 213 bp. Relative abundance of all T-RFs in electropherograms were between 1.5 to 75 %. *Npr* T-RFLP patterns were reveled two T-RFs: 102 and 213 bp, which were always present among different soil samples (Figure 17).

Decrease of *npr* specific T-RF numbers with increasing depth were observed and the highest number of *npr* T-RFs were found in upper soils. During the whole period of investigation higher number of *npr* T-RFs was detected at GC, LC and CP whereas the lowest values were measured at GK. Temporal fluctuation was not consistent at different sites: at CP maxima of *npr* specific T-RF numbers was noticed in April and at GC, LC and GK in July. In October the number of *npr* T-RFs was visible lower than in April or in July (Figure 16).

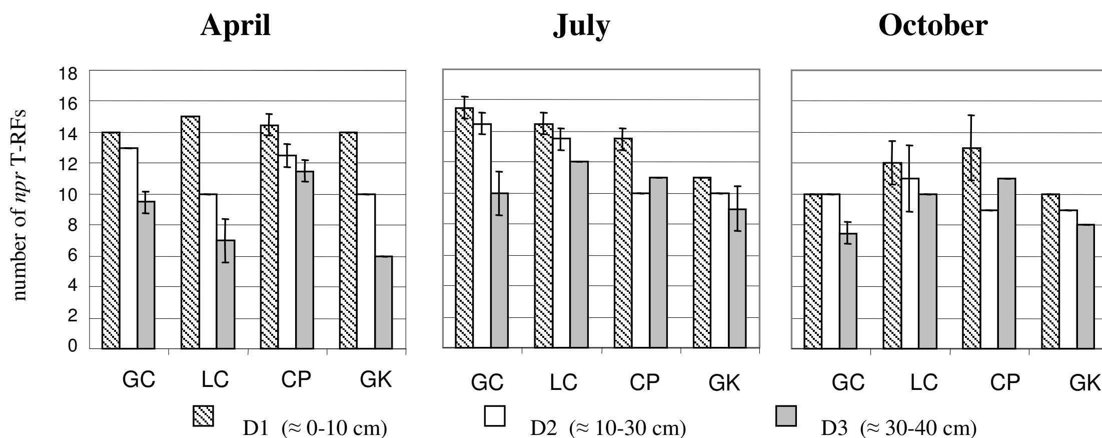


Figure 16: Number of *npr* T-RFs investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard deviation of means.

The dominating *npr* T-RFs in the samples collected in April were 102, 75 and 77 bp ranging from 20 to 73% of summarized heights of all peaks in electropherogram. In July, peak 102 bp was dominant ranging between 16 to 60% for most of the samples. In October T-RF 102 was exclusively the dominant one in all of the investigated samples ranging from 36 to 66% of summarized heights of all peaks in the electropherogram (Figure 17). However in the surface soil, dominance of one member (T-RFs) were less pronounce than in respective subsurface one (Figure 17, relative abundance).

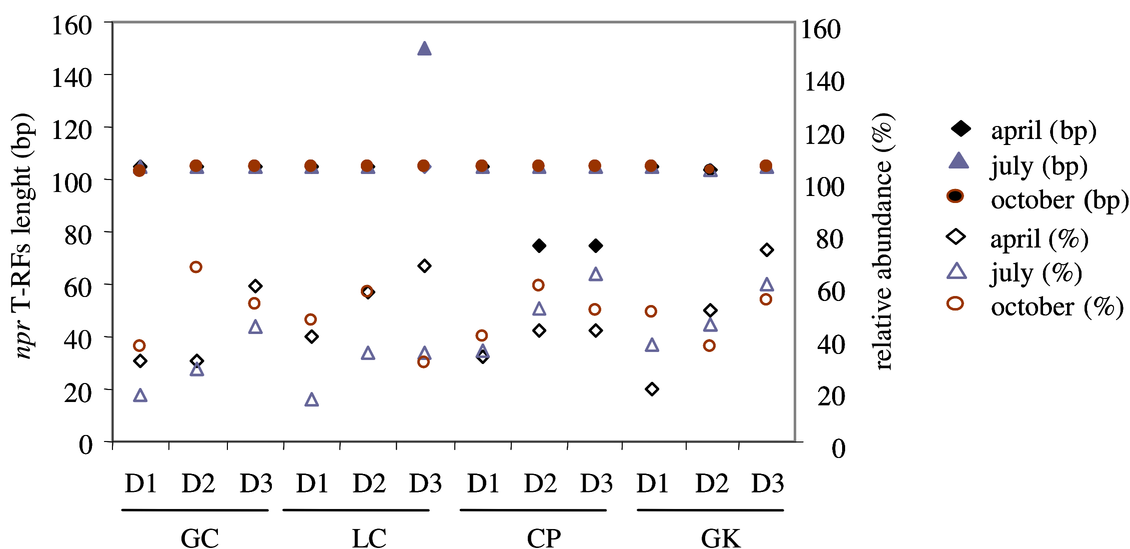


Figure 17: Length (bp) and relative abundance (%) of dominated *npr* T-RFs in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003.

In this study, PCA analysis proved to be a powerful method to discriminate between the three investigated parameters (site, depth and time). PCA ordering data are shown in Figure 18. PCA axis 1 explained 15.7% of the variability and PCA axis 2 explained 14.1%, with a cumulative percentage of 29.8%. The other PCA axis explained evidently less of variability (<9%). Effects of time, site and depth on components in PCA were tested using ANOVA analysis revealing significant effect of time on PCA axis 2 (P <0.002) due to the significant differences in community structure in July in comparison to April and October that displayed similar values. When the effects of site and depth on components in PCA were closely investigated for each of the month separately it was observed that depth tend to effect community structure in April and October and site-specific properties tend to influence the composition of *npr* communities in July. However observed effects were very weak (P=0.05).

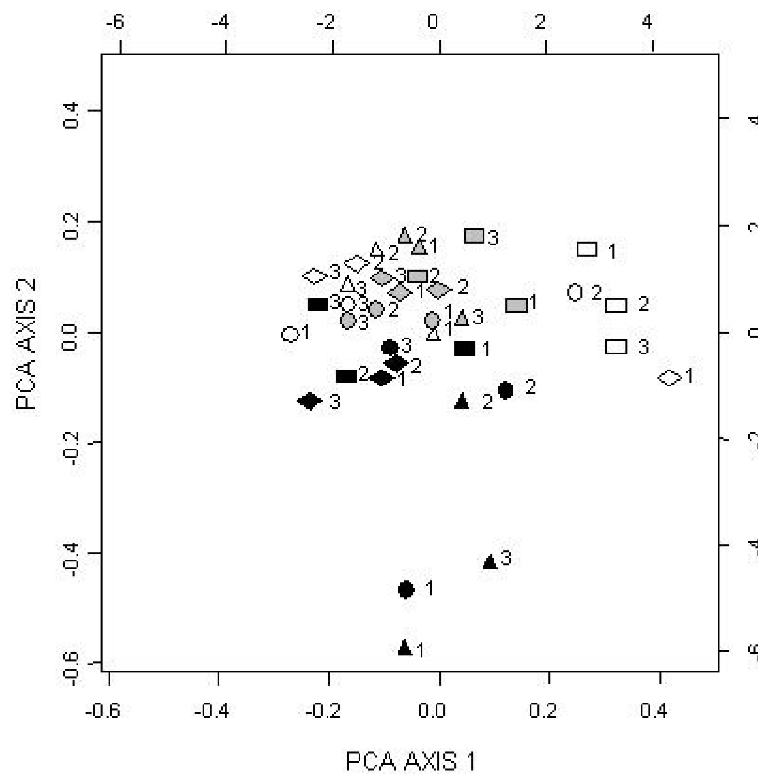


Figure 18: PCA ordering of data (axis 1 und 2) generated from *npr* T-RFLP fragments amplified from soil samples collected at field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. A symbol represents median of three measurements. Circles present GC, triangles LC, rectangles CP and diamonds GK. Gray represents April, black July and white October. The numbers are as followed: 1 for D1, 2 for D2 and 3 for D3.

3.7.2.2 *Sub* T-RFLP evaluation

Analysis of *sub* coding bacterial community of the soil samples provided 6-18 *sub* T-RFs that heights were above 1% of the summarized peaks heights present in the electropherogram. The different soil samples varied not only in the species richness that was estimated by determining the number of unique T-RFs but also in the numerically dominant T-RFs. *Sub* T-RF lengths ranged from 58 to 295 bp. Relative abundance of T-RFs in electropherograms were between 1.2 to 55%. *Sub* T-RFLP patterns revealed three T-RFs: 96, 125 and 130 bp that were always present among different soil samples (Figure 20).

The highest number of *sub* T-RFs was found in upper soils. Decrease of *sub* specific T-RF numbers with increasing depth was observed during the whole investigation period. Site-specific effect was also observed; the highest number of unique *sub* T-RFs was noticed at GK and the lowest at GC. The number of *sub* T-RFs did not considerably differ during the season. However tendency, although site or depth dependent, could be observed. At GC and CP the highest diversity was obtained in April whereas at LC the maximum values were noticed in July. At GK highest number of *sub* T-RFs was revealed in July and October (Figure 19).

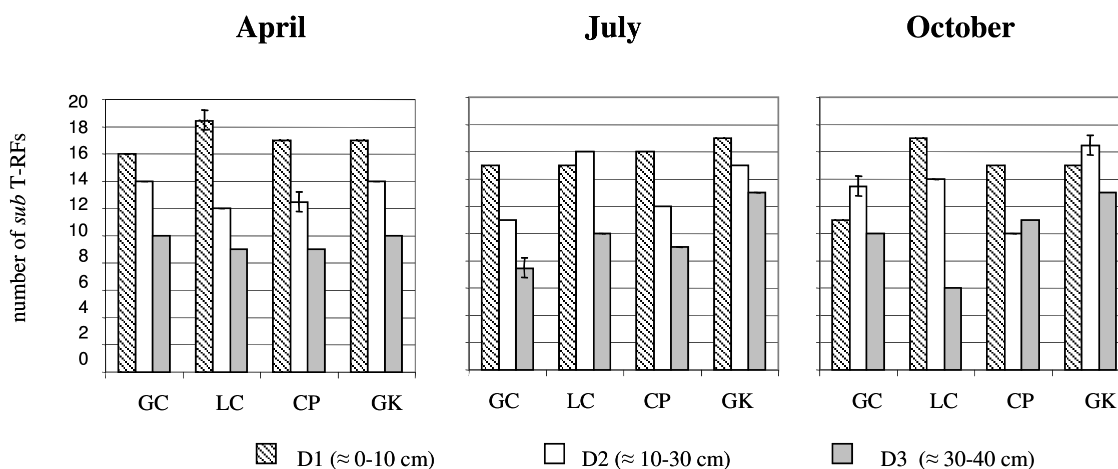


Figure 19: Number of *sub* T-RFs investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard deviation of means.

The dominated *sub* T-RFs in samples collected during the season at investigated field were 88, 96, 125 and 130 bp, which abundance ranging from 11 to 55% in a respective electropherogram (Figure 20). However surface and subsurface samples differ comparing the

dominance of T-RFs that was much less pronounced in the topsoils than in respective subsoils (Figure 20, relative abundance).

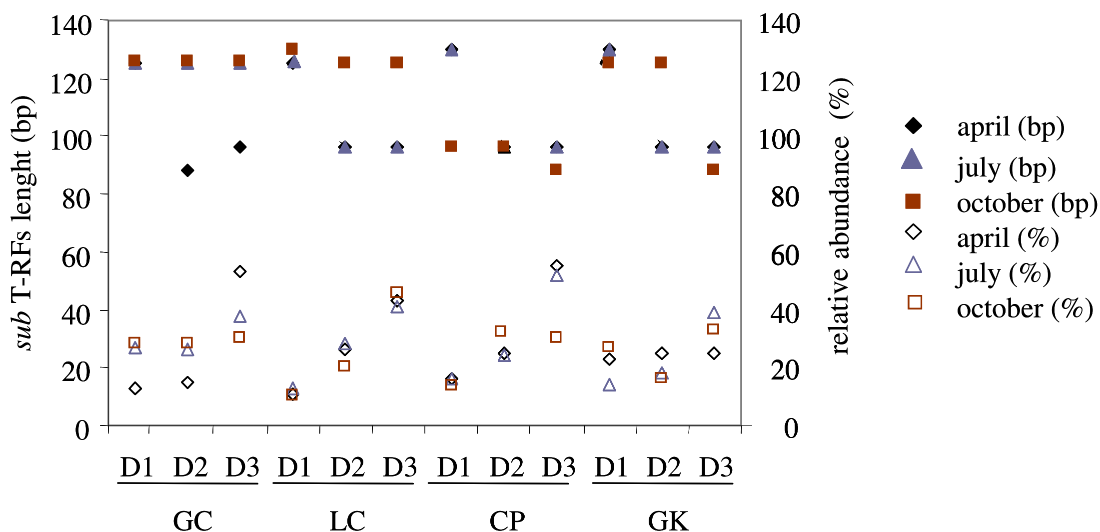


Figure 20: Length (bp) and relative abundance (%) of dominated *sub* T-RFs in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003.

To evaluate influence of the three investigated parameters (site, depth and time) on the *sub* community structure, PCA approach was applied. PCA ordering data are shown in Figure 21. PCA axis 1 explained 17.8% of the variability and PCA axis 2 explained 13.1%, with a cumulative percentage of 30.9%. The other PCA axis explained evidently less of variability (<9%). Effects of time, site and depth on components in PCA were tested using ANOVA analysis revealing significant effects of depth on PCA axis 1 (P <0.001) and effects of site specific properties on PCA axis 2 (P <0.001). When the effects of site and depth on components in PCA were closely investigated for each of the month separately, it was observed that depth effected community structure in April and July (P <0.001) and site-specific properties influenced the composition of *sub* communities in October (P <0.001).

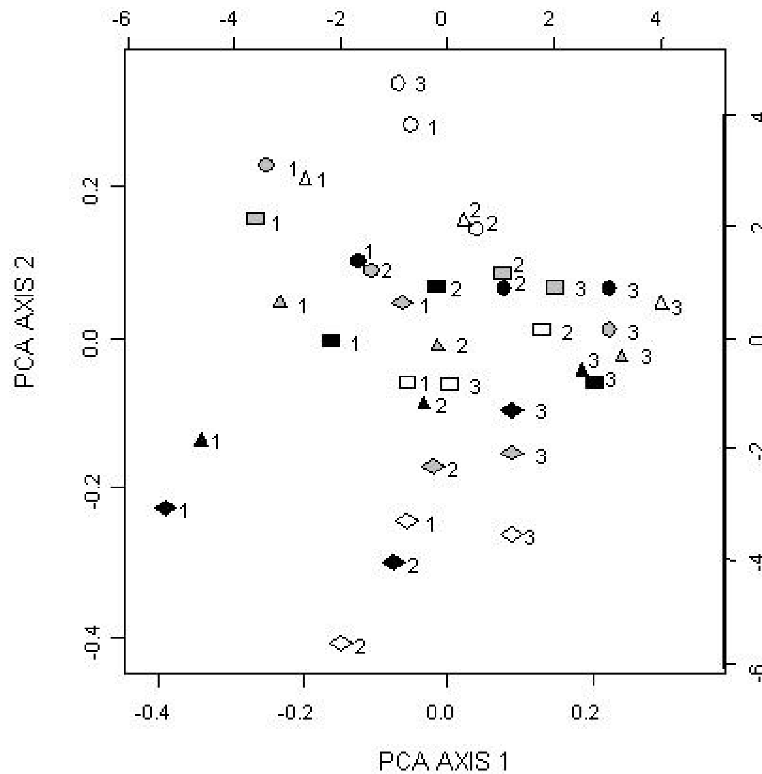


Figure 21: PCA ordering of data (axis 1 und 2) generated from *sub* T-RFLP fragments amplified from soil samples collected at field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. A symbol represents median of three measurements. Circles present GC, triangles LC, rectangles CP and diamonds GK. Gray represents April, black July and white October. The numbers are as followed: 1 for D1, 2 for D2 and 3 for D3.

3.8 Establishment of *npr* gene bank

Four clone libraries were established based on the results obtained by T-RFLP analysis. FP*nprI* and RP*nprIIb* primers were used to obtain *nprB* PCR products from the top soils (D1) collected at four different soil types (GC, LC, CP and GK) in July. PCR products were purified and ligated into pCR®2.1 vector and transformed into chemically competent One Shot® TOP10 cells. 80 transformants per each of the four soil samples was randomly picked up and subjected to low yield plasmid isolation. After *Eco*RI endonuclease digestion positive plasmids that contain 503 bp long inserts (483 bp inserts +20 bp vectors) were sequenced to check for the specificity. 45 inserts were sequenced per clone library corresponding to 56% of positive clones.

Obtained nucleic acid sequences were translated to amino acids using ExPASy and compared to sequences deposit in GenBank databases using NCBI BLAST. Sequences that were not characterized as neutral metalloproteases were excluded from further analysis. The portion of non-proteases was in average 20%. The remaining sequences did not match any sequences in the databases. Finally only sequences of high quality were aligned to *npr* sequences of reference organisms beeing the basis for phylogenetic analysis. 94 sequences obtained from the four clone libraries were aligned as well as 12 reference sequences. These sequences were characterized with the gene bank name obtained from the four different soil types (GC, LC, CP and GK) and a number that marks the clone number within the clone library.

Three different treeing methods, distance matrix data with neighbor joining, parsimony and maximum-likelihood were applied for the interference of the phylogenetic tree. The obtained trees were visually compared to confirm the stability of the tree topology. Bootstrap analysis was used to estimate the reliability of the phylogenetic reconstruction (1000 replicates). The consensus tree established from the 1000 replicates is showed in Figure 22. A single sequence was chosen to represent all sequences in cluster consisting of >98% identical sequences. Based on their nucleic acid sequences the detected terrestrial *npr* sequences showed similarities from 33 to 100%. From the 120 sequences 24 exhibit 100% similarity.

The obtained soil *npr* sequences were highly diverse and spread through the whole phylogenetic group of *npr* sequences. In additions two monophyletic clusters could be seen that contain only *npr* sequences from the soil samples (I. and B). All obtained *npr* sequences

were only distantly related to already publish *npr* sequences. No *npr* clone was identical to already known *npr* sequences. In Cluster I. there were 50 *npr* sequences from the four different soil types. 17 were from the GC, 23 from CP, 9 from GK and only one from LC. Cluster A contained one GK, one CP and one GC sequence as well as *npr* sequence from *Vibrio vulnificus*. Cluster B was exclusively built from soil *npr* sequences from the four different soil types. A large part of them, 18 were representatives from LC, 9 were from GC, 6 from GK and only 1 from CP. Cluster C was built from one soil sequence from LC and three *npr* sequences from *B. vietnamensis, Paenibacillus polymyxa* and *Clostridium acetobutylicum*. Cluster D contains 26 soil *npr* sequences from three different soil types which were clustering with the *npr* sequence from *Thermoactinomyces sp. Npr* sequences originating from GC were not present in cluster D. LC was represented with 11, GK with 13 and CP only with 2 *npr* sequences. Cluster E comprised *npr* sequences from the cultured bacteria *Alicyclobacillus acidocaldarius, B. stearothermophilus, B. caldolyticus, B. megaterium, B. thuringiensis, B. cereus, B. anthracis* and one *npr* sequence from GC.

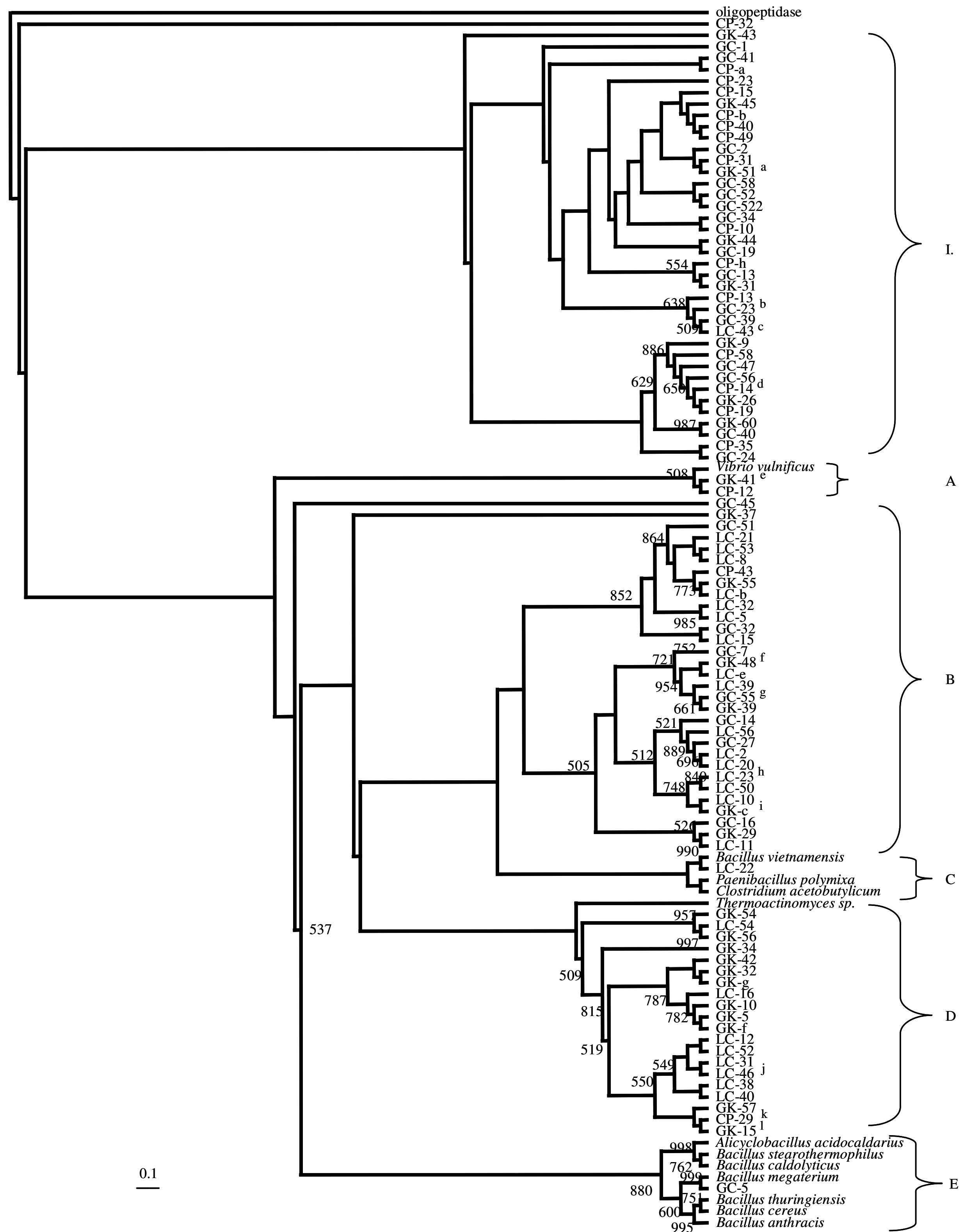


Figure 22: Neighbor-joining analysis of *npr* sequences based on 120 amino acids cloned from the soil samples. Clones obtained from this investigation are shown with soil type labeling and the number of the clone under respective soil type library. Bootstrap values for a total of 1000 replicates are given, values below 500 are omitted. A single sequence was taken to represent all sequences in cluster consisting of > 98% identical sequences: [a] GK-59, CP-3, 7, 34, 54, f; [b] GC-29; [c] GC-35, CP-2, 56; [d] CP- 6, 9; [e] GC-8; [f] GC-26; [g] GC-57, GK-49; [h] LC-30; [i] LC-3; [j] LC-16, 28, 36; [k] GK-58, CP-39; [l] GK-16.

3.8.1 Rarefaction analysis

Rarefaction analysis is used to estimate the diversity if sample size had been lower than it actually was and to assess the number of clones sufficient enough to encompass the bacterial diversity. Rarefaction curves were generated using the software "Analytic Rarefaction" (http://www.uga.edu/~strata/software/Software.html). The program uses the rarefaction equations for the expected number of species (E) given by Hurlbert (1971) and for the variance of the expected number of species (Var) given by Heck et al. (1975). Rarefaction's curves were made for the four *npr* clone libraries (n=120) and for each of clone library (n=30) separately (Figures 23 and 24). In this curves the number of different OTUs (operational taxonomic unit) is plotted versus the number of analyzed clones. OTUs were generated from distance matrix data where >98% similarity between *npr* sequences was criteria for OTUs grouping. As demonstrated in Figure 23 and 24 all curves, expect GC showed gentle slope. It indicates that the number of investigated samples relatively good cover the terrestrial diversity of *npr* sequences in respective soil samples. Further investigation of *npr* sequences from those samples would not drastically change obtain results. However, a plateau is not reached for GC curve, which exhibit steep slope. As indicates in Figure 23 and 24 four different libraries shared the same OTUs that often were not found more that once in one clone library (n=30), such terrestrial *npr* diversity was relatively good gain able by constructed four *npr* gene banks (n=120).

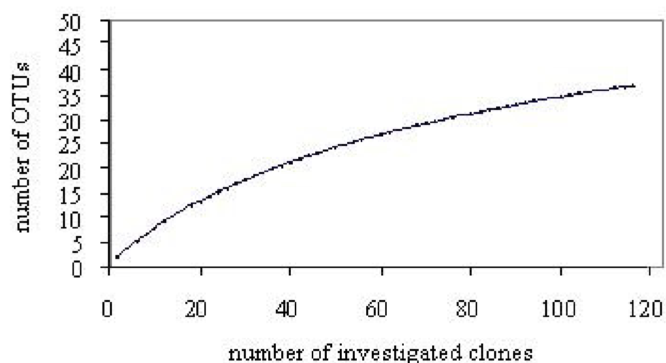


Figure 23: Rarefaction analysis of the 120 *npr* clones in soil samples from field site with four different soil types (GC, LC, CP, and GK) in July 2003. A calculation was preformed at the species level (> 98% sequence similarity).

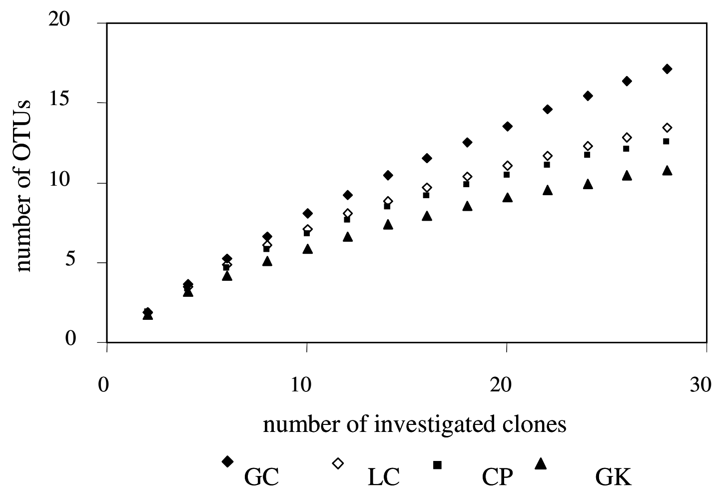


Figure 24: Rarefaction analyses of the *npr* clones in soil samples from field site with four different soil types (GC, LC, CP, and GK) in July 2003. Calculations were preformed at the species level (> 98% sequence similarity).

3.8.2 Specificity of the *npr* primer set used for phylogenetic analysis

To verify that the specificity of the new primer set matches that one used for T-RFLP a virtual PCR using the ModelInspector tool of the Genomatix software package (www.genomatix.de) was performed. Of the 61 *npr* sequences of different bacteria used for primer design, 22 could be detected allowing 2 mismatch in primer binding, thus showing nearly exactly the same specificity as the primer pair designed by Bach *et al.* (2001) which detected 21 sequences. Additionally a virtual PCR was performed against the data set of Bacteria (GenBank Release 146), again allowing two mismatches. 169 matches could be detected and 160 of them showed the expected size and specificity. 9 hits showed a wrong specificity, but the theoretical amplification products ranged with one exception from 660 to 1596 bp and could therefore be electrophoretically separated from products with the desired size and specificity. The constructed oligonucleotide sequence (RP*nprIIb*) and primer set used for T-RFLP analysis (FP*nprI*/RP*nprIIa*) is shown in Table 4 (see section 2.3.4).

3.8.3 Connection of cloned *npr* sequences to T- RFLP

To simulate the T-RFLP migration behavior of the obtained cloned *npr* sequences, they were cut to get 233 bp long fragments that had been used for the *nprA* T-RFLP assay. In a computer simulation using NEBcutter, all cloned *npr* sequences were cleaved with the same restriction endonucleases used to cleave the *nprA* PCR products amplified from the soil DNA. All sequences that showed >98% of similarity were investigated only once. The lengths of these hypothetical T-RFs were calculated and thus cloned sequences were associated to peaks in the electropherograms, considering that a difference of ±2 bp in the length of a T-RF is possible to occur due to the nature of the gel separation. 13 peaks (73%) found in the theoretical T-RFLP patterns obtained of the cloned *npr* sequences could also been found in the electropherograms. 6 theoretical T-RFs could not be found in the obtained electropherograms due to their small size (<50), which can not be detected by the used automated DNA sequencer. The not detectable theoretical T-RFs contributed to 27% of all cloned *npr* sequences.

4. Discussion

The aim of the present study was to investigate indigenous proteolytic bacterial community of field soils with focus on factors influencing their size, structure and activity. Furthermore, as physicochemical conditions of the surrounding environment exhibit a direct influence on the dynamics of extracellular protease activity, they were taken into particular consideration.

The second aim was to compare the structural and functional diversity of native proteolytic community in the soil system under multifactorial influences with attention to events deeper in the soil profile. To avoid biases recognized in culture-depended techniques, molecular biological tools were applied, which greatly afforded the study of proteolytic community in natural environments. The results should be interesting not only for ecologist but also for agriculture, industry and society as proteolytic bacteria community greatly influenced the level of available nitrogen in soil and with it contributing to soil functioning, health and plant productivity.

4.1 Usefulness of the primers used for analysis of total and proteolytic communities in soil samples

PCR primer sets for the detection of 16S rRNA, *sub* and *npr* genes has been successfully applied to study bacteria in bulk soil (Bach *et al.*, 2001) or in the rhizospheres (Sharma, 2003). 16S rRNA has been commonly used to study the bacterial diversity and abundance in various environments. 16S rRNA covers the phylogeny of the complete range of bacteria and it has essential function in all bacteria.

Npr is a marker for gene encoding for zinc depended neutral metalloprotease. *Npr* sequences are found in Gram-positive and -negative bacteria. However it is not common for all bacteria and only bacteria with proteolytic potential preserve the gene coding for extra cellular neutral metalloprotease. *Sub* is used as a marker for gene encoding for subtilisin. It is found among Gram-positive *Bacillus* ssp. species.

Universal bacterial primers FP16S rRNA/RP16S rRNA as well protease gene specific primers FP*nprI*/RP*nprIIa* and FP*subIa/*RP*subII* have been commonly used for bacterial community analysis. The used primer pairs are well established in literature (Bach *et al.*, 2002).

In order to amplify longer *npr* fragments the new reverse primer (RP*nprIIb*) had been constructed. Primer pair FP*nprI*/RP*nprIIa* showed almost the same specificity as FP*nprI*/RP*nprIIb*. The construction of new reverse primer was necessary in order to reveal much phylogenetic information about the diversity of *npr* coding bacterial community in the soil system.

4.2 Efficiency of soil DNA extraction

One of the aims of present study was to determine gene copy numbers as a determination of gene frequencies. Since DNA recovery significantly vary among different habitats or even soil types it was necessary to determine recovery efficiency of DNA extraction. Environmental DNA extracted from four different soil types and three different depths in October was added to 0.5 g of respective soil in a concentration providing ten times higher *npr* gene copies number than in unspiked soil samples. Recoveries of *npr* gene when directly added as environmental DNA to soil samples were between 36 to 62% for different soil samples. Since the extraction efficiencies varied among different soil samples and was <100%, the gene copies number obtained by Real-Time PCR was multiplied by a factor needed to reach 100% extraction efficiency in order to make comparative analysis of communities' size among different soil samples.

Several other studies have also determined DNA recovery efficiencies for soil or sediment samples. Recoveries varied in a range from 2.7 to 90% based on different extraction approaches and method used to estimate recovery efficiency. Mumy *at al.* (2004) added target DNA to sediment as whole cells or as purified plasmid DNA prior to extraction. They tasted three commercially available kits and found that recoveries ranged from undetectable to 43.3%. The FastDNA SPIN kit yielded the highest recoveries ranging from 19.4 to 43.3%. Their results also suggested that target DNA was recovered with similar efficiency when added as whole cells or as purified DNA.

The alternative to avoid problems connected to DNA recovery and extraction efficiency to Real-Time PCR was to quote protease gene abundance data to 16S rRNA gene considering that recovery efficiency for each of the target gene present in a soil sample is identical. However that approach is tried to be avoid since the variation in the number of gene copies

per genome as well as organism's growth strategy can drastically varied as would be discussed afterwards.

4.3 Number and composition of 16S rRNA, *sub* and *npr* genes per genome

DNA-molecular based methods were used to investigate dynamic of bacterial communities in soil. However it has to be considered that the number and composition of 16S rRNA, *npr* and *sub* genes per genome in prokaryotic organisms can differ such influencing results obtained by Real-Time PCR, T-RFLP and cloning analysis. Number of 16S rRNA genes per cell could be between one (many organisms) and 15 (*Clostridium paradoxum*) (Klappenbach *et al.*, 2001). High variation of *npr* copy number is also noticed. According to the NCBI databases (http://www.ncbi.nlm.nih.gov/) the number of *npr* copies can vary from one (many organisms) to seven (*Bacillus cereus* E33L). Genes encoding for neutral metalloproteases can occur on chromosome and plasmid. From seven copies of gene encoding for extracellular metalloproteases, five are chromosomal coding and two are plasmid coding occurring on pZK467 plasmid as is known from *Bacillus cereus* E33L genome analysis (http://www.ncbi.nlm.nih.gov/). The lowest variability is connected to *sub* copy number. Most proteolytic bacteria possess one copy of *sub* per cell or in some cases two copies of *sub* gene occur per genome (*Bacillus amyloliquefaciens* ATCC 23844).

Klappenbach *et al.* (2000) also found that the number of rRNA genes correlates with the rate at which phylogenetically diverse bacteria respond to resource availability. Soil bacteria that formed colonies rapidly upon exposure to a nutritionally complex medium contained an average of 5.5 copies of the small subunit rRNA gene (r-strategists), whereas bacteria that responded slowly contained an average of 1.4 copies (K-strategists).

It is important to notice that the microorganisms may exhibit r- or K-selected growth strategy or biphasic growth in response to environmental changes, showing r-behavior under the first set of conditions and K-behavior under a second. The r-selection manly favors population growth and it is characterized by higher number of gene copies per genome in comparison to K-selection. Although different values of qCO2 and SIR/CFE ratio indicate that the microbial community in the crop field dominated by r-selected species, seasonal, spatial and vertical differences as a result of differences in resource availability and quality are not surprising and are noticed.

Surface soils are predominantly characterized as nutrient rich and deeper soils as nutrient limited environments. Kristufek *at al.* (2005) were investigated the growth strategy of heterotrophic bacterial population along four success ional stages on spoil of brown coal colliery substrate in the surface (0-50 mm) and mineral layer (100-150 mm). They found that surface displayed a trend indicative of an r-K continuum in contrast to subsurface, where an r-strategy persisted. In contrast, it has been demonstrated that surface soil (0-5 cm) habiting fast growing, zymogenous portion of the microbial community opposite to the community habiting the deeper horizons (Fierer *et al.*, 2003). In addition, investigation of seasonal differences of microorganism's growth strategy revealed that percentage of fast-growing cells varied during the season and was significantly correlated with quantity and quality of available substrate (Smit *et al.*, 2001).

Since the number of gene copies per genome varies in different prokaryotic organisms and it is depended on microorganisms physiological status and growth strategy only assumption could be made about the size of total or proteolytic bacterial population. Changes in cell physiology resulting from changes in environmental conditions might influence variation of the copy number of particular gene pro genome. Thus, intrastrain (intraspecies) variability and the presence of multiple copies of particular gene might influenced results obtained from the complex soil ecosystem and therefore should be analyzed with caution. These findings have to be taken into account when the size or even diversity of bacterial population especially among soil profiles or among sites is compared.

4.4 Factors affecting abundance of 16S rRNA and protease genes in soil

Microbial communities respond rapidly to changes in their habitat. There are many abiotic and biotic environmental factors that induce shifts in microbial communities. These shifts appear to be due to interaction of different factors at the same time making our understanding of regulatory mechanisms of microbial communities more difficult.

Quantitative PCR approach revealed that vertical, site and temporal affects influence the abundance of 16S rRNA and bacterial proteases genes. The vertical and site effects were more underlined than temporal impact although time effect was not neglect able at all. The source availability might be beside the soil texture the main factor responsible for the observed variation in microbial community size.

The effect of soil depth on the size of proteolytic communities was until now only studied by Bach & Munch (2000). They found significant differences between top- and subsoil in MPN count of proteolytic bacteria in March and no differences in October at the same arable field. Only slight decrease of proteolytic bacteria number by increasing depth could be observed in other terrestrial ecosystems (Bach & Munch, 2000). Many studies have shown that total microbial biomass was significantly affected by depth (Agnelli *et al.*, 2004; Blume *et al.*, 2002; Fierer *et al.*, 2003; Taylor *et al.*, 2002) being considerably lower in subsurface environment as well as the amount and heterogeneity of available substrate (Agnelli *et al.*, 2002; Boehme *et al.*, 2005; Lorenz & Kandeler, 2005; Marx *et al.*, 2005; Taylor *et al.*, 2002).

The decline in the size of 16S rRNA and bacterial protease genes through the soil profile observed in the present study is predominantly a function of both, decreasing amount and quality of organic matter content. In accordance are the measurements of nutrient contents (WEOC, WEON, organic C, and total N) that exhibit the same pattern as investigated microbial properties. Especially carbon sources seem to have a significant influence on observed changes in bacterial community size. According to is the observation of total N and WEON contents that correlated positively with the bacterial number but with lower correlation coefficients for each property compared to organic C and WEOC. Many studies have demonstrated that carbon availability is the main factor limiting bacterial growth in agricultural soils (Alden *et al.*, 2001) although nitrogen has been of the same importance too (Christensen *et al.*, 1996; Duah-Yentumi *et al.*, 1998; Paul & Clark, 1996).

Moreover, differences in water content within studied horizons were low and only a slight increase in deepest soils could be seen in comparison to the upper ones. The absence of the correlation between water content and the number of respective genes seems to indicate that it is not a factor influencing variability of bacterial population within the investigated profiles. Although no relationship could be found between soil water content and soil bacterial size, Van Gestel *at al.* (1992) reported about positive relation between microbial biomass and soil moisture. Consequently, the substrate availability might be the main factor affecting abundance of soil bacterial communities among the investigated soil profiles.

The effect of source availability had been also seen when comparing size of proteolytic communities among different sites. The amount of organic C and total N was lowest at LC

where also the lowest abundance of proteolytic genes were measured and highest at GK where the highest number of protease genes were found, supporting the hypothesis that the carbon and nitrogen amount influence the size and activities of proteolytic microbes, which was also postulated by Ladd and Butler (1972).

Additional factors as soil texture seem to regulate bacterial properties at the investigated study site. LC was characterized by lowest clay/silt and high sand content in comparison to CP and GK. Soil texture is an important parameter that influences many soil characteristics such as organic matter storage, water-holding capacity and soil architecture, which in turn greatly effect soil biological properties as microbial biomass, primary production, enzyme activity and biodiversity. Therefore different soils with respect to soil structure might have different capacity to preserve necessary substrate heterogeneity and quantity and to facilitate spatial isolation within the soil matrix (Scott Bechtold & Naiman, 2006) .

Numerous of studies have been demonstrated that the bacterial number and biomass are concentrated in clay or silk particles (Emmerling *et al.*, 2001; Joergensen & Castillo, 2001; Zeller *et al.*, 2001). Although Ladd & Butler (1972) reported about high correlation between clay contents and proteolytic activity, nothing has been known about the influence of soil texture on the abundance of soil proteolytic bacteria.

From the other investigations it is known that activities of bacterial enzymes are highest in the soil small particles fraction (Marx *et al.*, 2005) whereas activity of xylanase as fungal indicator was higher in sandy fraction. Using PLFA and 16S rRNA-DGGE analysis Kandeler *et al.* (2000) confirmed those findings demonstrated higher bacterial biomass in silk and clay particles whereas in sand fraction fungal community dominated. One of the possible explanations is protection of microbes from predators within silk and clay particles. Additional, sand fraction is known as nutrient poor environment where consequently high level of bacterial and fungal competition occurs (Sessitsch *et al.*, 2001).

pH variations between the investigated sites were very low and there were no differences in pH between LC and GK. The highest 16S rRNA gene copy number was noticed at CP with lowest pH suggesting that pH was not an important factor regulating the size of bacterial communities among different sites. It is suggested that microbial biomass decreases only with distinctive pH changes like from 4.8 to 2.9 whereas changes from 6.8 to 7.7 had no effect on

microbial biomass (Lorenz & Kandeler, 2005). Therefore substrate availability could be beside soil texture the main factor influencing size of protelytic bacterial communities within different sites in the present study.

Seasonal changes in abundance of microbial communities are due to environmental conditions, temperature and soil moisture as well as plant root growth or plant residues during the growth and after harvesting. Many studies that have been investigated influence of seasonal variation on microbial community's size. Very often identified trends are contradictory (Wardle, 1998). However, it is extremely difficult to interpret seasonal differences against a background of considerable spatial variation.

In the present study, temporal changes in bacterial protease gene number were not consistent for the different sites or depths indicating that the impact of seasonal variation on the size of bacterial communities depended upon the sampling site or depth. However general tendency in the *npr* copies number was observed. *Npr* copies number exhibited the highest values in spring (exception LC) and decreased until autumn. The *npr* coding population seems to be positively affected by higher moisture, lower temperature and fertilizer application. It has been reported that applications of mineral fertilizer increase soil organic matter (SOM) quality and quantity (Galantini & Rosell, 2005) leading to enhanced bacterial growth. Addition of mineral nutrients may accelerate decomposition of nutrient limited substrates (Newell *et al.*, 1996) and enhance decomposition rates (Neher *et al.*, 2003). These findings might explain increased number of gene encoding for neutral metalloprotease in April.

No general tendency could be observed for the seasonal dynamics of 16S rRNA or *sub* genes, indicating that the dynamics of the 16S rRNA and *sub* coding population was more influenced by soil specific properties than by seasonal induced changes.

One of the aims of present study was to investigate relationship between protease and 16S rRNA gene numbers indicating that changes in abundance of protease genes are due to the changes in 16S rRNA gene pool size. Positive significant linear relationships between *sub* or *npr* and 16S rRNA copy number during the whole investigation period were observed, however relation was stronger for *sub*/16S rRNA than for *npr*/16S rRNA genes. The abundance of protease genes were presumably regulated by total community size but also factors affecting differently total community size and size of specific groups within total

bacterial community. However, it is very difficult to scale protease gene abundance data to 16S rRNA gene abundance since only assumption can be made about the number of copies of particular gene per genome in soil system as is discussed above. Although the number of gene copies is suspected to be stabile under stabile environmental conditions we do not know about the ratio between protease and 16S rRNA gene copies numbers, especially among different soil depths.

4.5 Factors affecting potential proteolytic activity in soil samples

Considerable spatial and seasonal variation in potential proteolytic activity was noticed in the present study. During the whole investigated period and at all sites potential proteolytic activity significantly decreased by increasing depth following the decrease of total and proteolytic bacteria as well as decline of nutrient contents. Bach & Munch (2000) also observed considerably lower proteolytic activity in subsoils than in corresponding topsoils. Generally lower rates of enzymatic activity in the subsoil has been observed in many studies (Lorenz & Kandeler, 2005; Taylor *et al.*, 2002).

Different protelytic potential within the profiles could mainly be explained by differences in the number of soil microbes and nutrient contents. Previous studies related increased proteolytic activity in soil to the enhanced growth of microbial communities (Asmar *et al.*, 1992; Asmar *et al.*, 1994; Holt & Mayer, 1998; Kandeler *et al.*, 1999; Watanabe & Hayano, 1995) or to changes in the nutrient status (Lorenz & Kandeler, 2005; Taylor *et al.*, 2002), or both. Opposite, Bach & Munch (2000) found no correlation between potential proteolytic activity and number of total or proteolytic bacteria in soil, however by estimating the number of proteolytic bacteria by culture dependent techniques.

A positive correlation between potential proteolytic activity and the number of 16S rRNA, protease gene copies number, organic C, total N, WEOC and WEON was found in this study. Therefore it has been suggested that the size of bacterial communities and especially the availability of potential substrate, which was concentrated in topsoil, could be the main factors that influence proteolytic activity among the soil profiles.

Beside the vertical effects potential proteolytic activity was also significantly influenced by site features such being constantly lower at LC and reaching the maximum at GK. Varying

activities between different sites may result from diverse soil properties, such as soil texture and nutrient contents as well as changing numbers of soil bacteria. As mentioned before LC was characterized by lowest clay/silt and extremely high sand contents in comparison to GK, which is in a accordance with a previous study by Marx *et al*. (2005) that found dominant proteolytic activity in the clay/silt fractions where also most of the amino acids are located. Ladd & Butler (1972) also reported about high correlation between clay contents and proteolytic activity.

The nutrient content was lower at LC and higher at GK. This findings correspond with the findings of Ladd & Butler (1972) who observed positive correlation between activity and organic matter content. Considering that proteinaceous material represent 40 % of total N in soil it has been suggested that proteolytic activity may be explained by availability of potential substrates (Marx *et al.*, 2005).

A seasonal trend in the potential proteolytic activity was also noticed. The highest activity was measured in October (exception CP) and minimum of activity was noticed in summer (July). Temporal changes of proteolytic activity may be attributed to environmental conditions, like temperature and soil moisture as well as root growth, presence of plant residues, fertilizer application and seasonal changes of soil microbial communities' size.

Higher proteolytic activity in April in comparison to July could be attributed to application of higher amounts of fertilizer in spring and decreased moisture in summertime. Schloter *et al.* (2003) observed that an application of mineral nitrogen fertilizer in spring resulted in a significant increase of proteolytic activity. Application of mineral fertilizer could accelerate decomposition rate (Newell *et al.*, 1996) and increase soil organic matter (SOM) particularly on lands where as a result of cultivation significant loss of SOM was observed (Galantini & Rosell, 2005). Therefore it is possible that increased SOM quality and quantity enhances microbial growth followed by depletion of nitrogen. Alden et al. (2001) reported that microbial growth is mainly limited by deficiency of C and after intensive increase of microbial biomass nitrogen limitation could be observed. This might explain higher proteolytic activity in this study one month after mineral fertilizer application.

Decrease of proteolytic activity in July was possibly due to reduction of soil moisture. It has been observed that a reduction of soil moisture decreased protease activity by 15–66% depending on annual period and soil depth (Sardans & Penuelas, 2005).

Potential proteolytic activity increased again after harvesting and accumulation of plant residues. This finding is in good agreement with previous studies that suggested that an addition of high energy material increases overall enzyme activity (Asmar *et al.*, 1994) thus the rise of proteolytic activity in this study in October is likely due to increased amount of organic substrates given by plant residues after harvest. Brosius *et al.* (1981) also postulated that increased enzymatic activity is induced by an increased amount of potential substrates.

Although proteolytic activity was positively correlated with the number of 16S rRNA and protease gene copy, correlation coefficients between activity and number of total bacteria were visibly lower compared to the number of proteolytic genes. It suggests that the potential for soil proteolytic activity is not overall distributed within the bacterial population but rather is a feature of one of its parts. These observations are complied with the findings that showed that *Bacillus* spp. are the major source of soil proteases (Watanabe & Hayano, 1994) and that proteolytic activity did not correlate with total but only with soil proteolytic bacteria (Watanabe & Hayano, 1995).

Hydrolysation of peptide compounds in soil are due to activity of extracellular proteases that are present in soil as free enzymes or are associated with clay minerals or humic colloids protecting them from the degradation which, despite affecting their catalytic potential, may facilitate enzyme activity to persist in soils.

Overall positive relationship between *sub* gene number and potential proteolytic activity or *npr* gene number and proteolytic activity was found. However when relationship between both measured parameters in two entirely diverse soil samples were closely investigated no relation could be found in soil sample with highest clay content or on the contrary the relationship was very high in the soil sample where the sand fraction dominated such indicated stabilization of proteases on clay particles and absence of the feature in sandy soil where the potential proteolytic activity may be mainly due to free enzymes in soil.

However this indication must be taken with caution because proteolytic activity in soil could be connected with other proteases except subtilisin and neutral metalloproteases or simply because no factor was limiting the activity. Also expressional level of genes coding for extracellular proteases might be quiet different in different surrounding environments.

4.6 Comparison of *npr* and *sub* gene communities by T-RFLP

T-RFLP allows the analysis of a large number of samples and it is a very good tool for initial screening of similarities and differences between different bacterial soil communities (Kirk *et al.*, 2004). Use of T-RF peak number as a measure of species richness in bacterial communities has been reported (Dunbar *et al.*, 2001). However an individual T-RF does not always represent an individual species or genus and different restriction enzymes can suggest different levels of phylotype richness within a community or different trends between communities. Despite the limitations in identifying individual species or in describing species richness, the T-RF method was useful for detecting compositional differences in complex soil communities.

In the present study, T-RFLP method was employed to investigate the vertical, site and temporal changes in bacterial community structure at the study site. This is also the first culture-independent analysis of the structure of indigenous proteolytic bacteria populations in environmental samples and the first application of T-RFLP technique to characterize composition and composition dynamic of the microbial gene pool with respect to bacterial proteolytic genes. PCR based methods have recognized biases when used in a multi template manner as is demanded for community analysis (Qiu *et al.*, 2001). However, relative comparison of microbial community structures have been routinely applied based on such methods overcoming biases recognized in culture-depended techniques. In general, carefully optimized, the T-RFLP has proven to be a reproducible, accurate and representative tool for community fingerprinting.

T-RFLP profiles of *npr* community revealed 7-15 T-RFs when labeled digested PCR products were separated, detected and analyzed, or 6-18 T-RFs when *sub* community was investigated. Manual check of electopherograms revealed no differences between the fingerprints of the replicates from the same soil sample or differences were very low. This suggests a low degree of variability caused by sampling, DNA extraction and T-RFLP analysis. There were

differences in the *npr* or *sub* community profiles between different soil samples with respect to the number and relative abundance of the peaks, each representing one sequence within the complex community, highlighting the presence or dominance of different bacteria in different soil samples. This allow for comparing proteolytic bacterial community structure and for assessing community dynamic.

4.6.1 Comparison of *npr* and *sub* T-RFLP profiles in soil samples

T-RFLP analysis approach revealed decreasing diversity of *npr* gene by increasing depth and a noticeable temporal variation in the *npr* bacterial community structure displayed by T-RFLP and ANOVA analysis. The possible reason for observed patterns in a *npr* community structure might be mainly quantity and heterogeneity of available substrate as well as spatial isolation caused by varying water amount and connectivity of soil particles.

The data suggested that the composition of proteolytic bacterial population in July differed from the population at the other dates. Also *npr* diversity was greater in July than in April or October. Such an exclusivity of July samples might be result of increasing level of quality and quantity of available substrate applied to the soil through fertilization, decaying plants' and fungal material left in the soil after fungicide application as well as from root exudates.

In a study of Galantini & Rosell (2005) fertilization has been shown to increase the level of crop residue production. Returning this residue to the soil, especially in reduce tillage systems has a positive impact on soil organic matter (SOM) quality and quantity. Plant root growth however has the potential to induce both stimulatory and negative effects based on the dual effects of plants on soil microbes. Plants could stimulate microorganisms through C addition or suppress them through resource competition (Wardle, 1998). In this study differences within the season in the examined system indicated stimulatory effects of plants exudation on *npr* coding bacterial community. In accordance are the results of the measurement of water extractable organic carbon (WEOC). WEOC was next April highest in July thus explaining an increased level of diversity for the summer sampling date.

Due to the access to those readily and easily available substrates microbial community might show the full potential in term of diverse community structure. Relatively high diversity in

July suggests that various different *npr* coding bacteria benefit from those available and heterogeneous substrates.

Although is known that soil is reservoir of extreme diversity, the present study confirmed that diversity of *npr* bearing bacteria is lower in deeper than in upper soils and is not spread equally through the 40 cm deep soil profile. Moreover, the dominance of one or few members was less pronounced in topsoils than in corresponding subsoils. This is in accordance with a study by Zhou *et al.* (2004), in which diversity of soil microbial community was determinated by SSU rRNA gene cloning approach and a much higher diversity in surface environment than in the corresponding subsurface one was detected as well as no dominant OTUs was noticed in topsoil quite opposite to subsoil. They hypothesized that spatial isolation and resource heterogeneity explain the differences in microbial diversity patterns observed among the soil profile. Because of usually low soil moisture in the surface, connectivity of soil particles decreases and a high level of spatial isolation occur allowing for maintenance of diverse types of microorganisms and an increased level of diversity. In contrast, in subsurface excess water facilitates a high level of connectivity. Consequently, part of diversity is lost by competitive exclusion due to the nutrients and microbes flow.

They also pointed to the essential role of carbon resources in regulating microbial diversity and composition. Greater substrate quantity and heterogeneity in upper layers facilitate high-level of microbial diversity, the trend that was also observed in the present study. Similarly, low level of diversity in deeper horizons has been noticed by Agnelli *et al.* (2004). They postulated that availability and nature of organic matter could be considered as the main factor in controlling diversity and structure of the microbial community among the soil profiles. These findings indicate that substrate heterogeneity and availability are beside spatial isolation a major factor in the selection of bacterial community.

T-RFLP analysis of *sub* community revealed decreasing diversity of *sub* genes by increasing depth as well as a noticeable vertical and site specific variation in the *sub* bacterial community structure displayed by T-RFLP and ANOVA analysis. The data suggested that the composition of *sub* proteolytic bacterial population at D1 or D2 considerably differed from the population at D3. Moreover composition and diversity of the *sub* community at GK differed from those at other sites (GC, LC and CP) exhibiting highest number of T-RFs at GK during the whole period of investigation. Sesonal changes did not considerably affect the

structure and diversity of *sub* specific community. Time effect was strongly dependend on soil vertical and site-specific properties.

As noticed for *npr* coding community, quantity and quality of available substrate as well as spatial isolation caused by varying water amount and connectivity of soil particles could be the main reason for observed patterns in a *sub* community structure among the investigated soil profiles. The measurements of nutrient contents (WEOC, WEON, organic C, and total N) revealed the highest concentration in the surface soils where also the highest diversity of *sub* coding community was observed. The observed patterns were in accordance with the study of Zhou *et al.* (2004), and Agnelli *et al*. (2004), which noticed decreased bacterial diversity in subsurface environment according to the decreased quantity and heterogeneity of available substrate. Moreover, the absence of dominance of one or few members in the topsoil noticed in the study of Zhou *et al.* (2004), was as well confirmed for *sub* diversity analysis.

Interesting, the number of *sub* T-RFs was highest at the GK where the lowest number of *npr* specific T-RFs was found. Sampling site GK contained the highest silt and considerably lowest sand content in comparison with the sites GC, LC and CP, thus possibly providing more convenient growth conditions for members of the bacteria bearing the gene coding for subtilisin than for the *npr* coding population. Soil physicochemical properties can be the powerful factor in a selection of inhabiting bacterial population such stimulating the growth of one population and repressing the other ones (Kandeler *et al.*, 2000).

4.7 Phylogenetic analysis of soil *npr* gene community

To confirm the specificity and to rule out possible limitations of T-RFLP analysis but also to obtain detailed information about the composition of the soil *npr* coding bacterial communities, *npr* PCR products were cloned and sequenced. Most of the fragments from the clone library (73%) were detected in the T-RFLP fingerprints from the complex proteolytic bacterial community. Twenty-seven percent of clone sequences revealed T-RFs with lengths less than 50 bp, which were therefore under detection limit. The revealed T-RFs were not essentially the most abundant ones. Observed discrepancies are presumably the result of biases introduced by each method separately thus differently affecting the results of molecular biological measures of diversity. Although is not to exclude that both methods, T-RFLP and cloning, introduce biases, Lueders and Friedrich (2000) concluded that the cloning step in particular was subject to bias. They observed large fluctuation in communities' composition of *Archaea* over time only in clone libraries. Despite the finding that clone abundance in the library does not necessary reflect the abundance of particular organism within complex bacterial communities (Head *et al.*, 1998), the present *npr* clone library revealed valuable data about the structure of a *npr* coding community in a soil system. It is also so far the first look at the actual (*in situ*) composition of *npr* coding proteolytic communities in the environmental samples.

Based on the T-RFLP results four clone libraries were established for the soil samples collected in July. The number of T-RFs and the composition of the *npr* community in July significantly differed from those investigated in April and October. T-RFLP analysis as well revealed that the number of T-RFs was considerably higher in surface than in respective deeper soils thus providing many diverse sequences for cloning analysis. In order to reduce the data set, only samples from upper soils were considered for the time consuming cloning procedure. For the cloning approach we constructed a new reverse primer to amplify a longer *npr* fragment. The specificity of both primer sets, used for T-RFLP or cloning analysis, were almost the same. The construction of new reverse primer was necessary in order to gain more phylogenetic information about the diversity of *npr* coding bacterial communities. The primer set used for T-RFLP analysis was successfully applied for the investigation of abundance of *npr* gene in soil system. However, it was difficult to gain relevant phylogenetic information with the short *npr* fragments.

Phylogenetic analysis of the soil *npr* sequences revealed that most clones show only poor homology to isolates previously obtained from various environments. Cloned *npr* sequences were highly diverse and spread through whole phylogenetic group of *npr* sequences. As far as is known there are no comparable data on phylogenetic analysis of soil *npr* sequences. Therefore comparative analysis of literature with data presented in this study was not possible. The *npr* community in surface soil at this study site included sequences related to *npr* sequences from *Vibrio* ssp., *Bacillus* ssp., *Thermoactinomyces* ssp., *Paenibacillus* ssp., *Clostridium* ssp. and *Alicylobacillus* ssp. However the largest proportion of cloned *npr* sequences, approximately 65% (cluster I. and B, Figure 22) did not branch with any known proteolytic bacteria, indicating that investigated soils contain uncharacterized *npr* proteolytic microbes.

A remarkable diversity of extracellular metalloprotease present in microbial communities was already noticed isolating *npr* coding bacteria. *Npr* coding genes have been found in Gram-positive and -negative bacteria as well as in both pathogenic and non-pathogenic species. Bacterial metalloproteases are known to be involved in pathogenesis and to serve useful function as degradation of inert polypeptide compounds in soil. Haese & Finkelstein (1993) postulated that ubiquity and conservation of extracelullar metalloproteases in the microbial world provide them survival advantages and they predicted that additional bacterial metalloproteases would be found in future. That was partially confirmed by results obtained in present study. However an evaluation of physiology and ecological role of *npr* in diverse soil bacteria was difficult because most of the detected *npr* sequences showed poor homology to already known ones.

Although *npr* sequences from the investigated soil samples were spread through the whole phylogenetic *npr* tree, the consensus tree revealed several important patterns. First, the vast majority of clones did not cluster with any known sequences of *npr* bearing proteolytic bacteria, indicating that the four soils contained hitherto uncharacterized *npr* bacteria. Second, the majority of the clone sequences from the investigated field formed a distinct cluster (I.) containing only soil *npr* sequences. Third, tree organization indicated that the four proteolytic *npr* soil communities are related with some overlap indicating that different soil properties caused a distinct shift in the *npr* communities rather than dramatic changes. However differences based on different site's properties were noticed. Cluster I. (Figure 22) contained only one sequence originating from the clone library of site LC while representatives of

libraries GC (17 clones) and CP (23 clones) were dominating. In contrast in cluster B a large number of clones obtained from sampling site LC (18 clones) were prevailing, and to a lesser extent clones from site GC, GK and CP (9, 6 and 1 respectively). In the *Thermoactinomyces* ssp. cluster there were no *npr* sequences from GC detected (Figure 22 Cluster D).

It is very difficult to explain observed patterns based only on already measured parameters. *Thermoactinomyces* ssp. is widely prevalent in agricultural soils thus one should not expect the exclusion of this group just from one of the investigated sites. Rarefaction analysis revealed that the clone libraries of site CP, LC and GK contain a sufficient number of clones to encompass the bacterial diversity. Increasing the number of clones would not significantly change the already observed *npr* diversity. On contrary, gene bank GC did not completely cover the diversity of *npr* sequences present at the investigated site, thus no reliable conclusions could be drawn about the presence of the actually dominant and representative *npr* population in the library established for site GC. The absence of *npr* sequences related to that of *Thermoactinomyces* ssp. from the GC library could be partly the result of a high level of diversity of *npr* sequences at site GC that were not completely covered by the respective gene bank.

Nevertheless what is captivating is the fact that the clone libraries of sites LC (11 clones) and GK (13 clones) had many representatives clustering to this group. However, only two sequence of site CP and no sequences of GC belong to this cluster. The reason for observed patterns could be also found in the physicochemical properties of the different sampling sites. Isolation procedures of members of the genus *Thermoactinomyces* from two different soils identified many members of this group with growth optima at pH 5 (Yallop *et al.*, 1997). Sampling sites LC and GK (D1) revealed a lower pH and lower clay content than sites GC and CP (Table 1), thus possibly providing convenient growth conditions for members of the genus *Thermoactinomyces*.

4.8 Conclusion and perspectives

Proteolytic activities are essential for nitrogen turnover processes in soil. Native proteolytic bacterial community plays a critical role in regulating the production of extracellular proteases such affecting the level of available nitrogen for plants and microbes in soil. Environmental conditions of surrounding habitats influence the properties of microbial community and are pivotal factor in controlling activity of once secreted extracellular enzymes. Therefore the effects of environmental conditions on proteolytic community size and composition as well as on activity of free proteases in soil may be determined in regulating the level of proteolysis. Although many studies have been investigated the influence of various environmental factors on proteolysis in soil systems almost all microbiological investigations are based on laboratory experiments and culture dependent analysis. Moreover, only limited research has been done to investigate dynamic of indigenous bacterial population coding for extracellular proteases. This reinforces the importance of present study, which deals with influence of naturally occurring and therefore varying environmental conditions on structure and function of native bacterial proteolytic community in soils.

Using different culture-independent techniques, Real-Time PCR, T-RFLP, cloning and casein measurement, marked differences of the size, diversity and activity of bacterial proteolytic genes community present in the different soil samples were shown. These differences can be related to differences in quantity and heterogeneity of available substrate, soil texture as well as spatial isolation caused by varying water amount and connectivity of soil particles. Mineral fertilizer application, soil cultivation and seasonal variation in temperature or soil moisture were the factors that additionally influenced the properties of bacterial proteolytic community in soil. The results of this study showed that the size, structure and activity of bacterial proteolytic community is highly dynamic; it showed remarkable spatial, vertical and seasonal variability due to the variation of the environmental conditions such greatly affecting the level of nitrogen turnover in soil. The molecular based analyses used in this study were valuable and sensitive tools to access the dynamic of complex proteolytic processes at the sites under investigation. For a detailed regulation of proteolytic activity a system biology approach is needed to understand the regulation of the induction of the corresponding operons in more details.

5. Summary

The aim of the present study was to increase our knowledge on proteolytic processes in agricultural soils. Diversity, abundance and activity of proteolytic populations were investigated with respect to spatial, temporal and vertical variability of naturally occurring environment factors. For this purpose an arable field in Southern Germany under integrated management was studied. The uniformly managed field shows pronounced soil heterogeneity with four different soil types. In April, July and October 2003 soil samples were taken from the four soil types at three different depths, up to 40 cm, according to the horizontation of the soil. A cultivation-independent approach was applied. PCR was employed to detect the genes encoding for neutral metalloprotease (*npr)* and subtilisin (*sub*). Abundance of genes of interest was estimated by the Real-Time PCR whereas potential proteolytic activity was measured by transformation of added substrate. The number and activity of respective genes showed high vertical, spatial and seasonal variation due to the variable physicochemical characteristics of investigated soils. Mainly factors as substrate availability and soil texture were responded for observed changes. Additionally, mineral fertilizer application, soil cultivation and seasonal variation in temperature or soil moisture influenced the size and activity of proteolytic community. First characterization of proteolytic bacterial community composition of the investigated soil samples was done by T-RFLP fingerprinting analysis. The structure of *npr* or *sub* communities was as well affected by vertical, spatial and seasonal variation. To obtain detailed information about the structure of the soil *npr* coding bacterial community, cloning and analyzing of *npr* PCR products were performed. Clone libraries were generated from the topsoil samples collected in July. Phylogenetic analysis of the soil *npr* sequences revealed that most clones have a poor homology to isolates previously obtained from various environments. Cloned *npr* sequences were highly diverse and spread through the whole phylogenetic group of *npr* sequences. The present study revealed a hitherto unknown diversity of proteolytic genes in the agricultural soil at the investigated study site. However, phylogenetic tree organization indicated that the four proteolytic *npr* soil communities are related with some overlap. Overall the results indicate considerable influence of soil physicochemical properties, management practice during the season and climatic seasonal changes on soil proteolytic community structure and function. Varied environmental conditions caused variability in the size, structure and activity of proteolytic community at the sites under investigation.

6. References

Agnelli, A., Celi, L., Corti, G., Degl'Innocenti, A. & Ugolini, F. (2002). The changes with depth of humic and fluvic acids extracted from the fine earth and rock fragments of a forest soil. *Soil Science Society of America Journal* **167**, 524-538.

Agnelli, A., Ascher, J., Corti, G., Ceccherini, M. T., Nannipieri, P. & Pietramellara, G. (2004). Distribution of microbial communities in a forest soil profile investigated by microbial biomass, soil respiration and DGGE of total and extracellular DNA. *Soil Biology and Biochemistry* **36**, 859-868.

Alden, L., Demoling, F. & Baath, E. (2001). Rapid method of determining factors limiting bacterial growth in soil. *Applied & Environmental Microbiology* **67**, 1830-1838.

Amann, R. I., Ludwig, W. & Schleifer, K. (1995). Phylogenetic identification and in situ detection of individual cells without cultivation. *Microbiological Reviews* **59**, 143-169.

Asmar, F., Eiland, F. & Nielsen, N. E. (1992). Interrelationship between extracellular-enzyme activity, ATP content, total counts of bacteria and CO2 evolution. *Biology and Fertility of Soils* **14**, 288-292.

Asmar, F., Eiland, F. & Nielsen, N. E. (1994). Effect of extracellular-enzyme activities on solubilization rate of soil organic nitrogen. *Biology and Fertility of Soils* **17**, 32-38.

Auerswald, K. & Kainz, M. (1990). Standortlicher Ueberblick ueber das Klostergut Scheyern. In: Auerswald, K. and M. Kainz (eds.). Forschungsverbund Agraroekosysteme Muenchen. Lehrstuhl für Bodenkunde der TU München, Weihenstephan, pp 41

Avrahami, S., Conrad, R. & Braker, G. (2002). Effect of soil ammonium concentration on N_2O release and on the community structure of ammonia oxidizers and denitrifiers. *Applied & Environmental Microbiology* **68**, 5685-5692.

Bach, H.-J. & Munch, J. C. (2000). Identification of bacterial sources of soil peptidases. *Biology and Fertility of Soils* **31**, 219-224.

Bach, H.-J., Hartmann, A., Schloter, M. & Munch, J. C. (2001). PCR primers and functional probes for amplification and detection of bacterial genes for extracellular peptidases in single strains and in soil. *Journal of Microbiological Methods* **44**, 173-182.

Bach, H.-J., Tomanova, J., Schloter, M. & Munch, J. C. (2002). Enumeration of total bacteria and bacteria with genes for proteolytic activity in pure cultures and in environmental samples by quantitative PCR mediated amplification. *Journal of Microbiological Methods* **49**, 235-245.

Badalucco, L., Kuikman, P. J. & Nannipieri, P. (1996). Protease and deaminase activities in wheat rhizosphere and their relation to bacterial and protozoan populations. *Biology and Fertility of Soils* **23**, 99-104.

Blume, E., Bischoff, M., Reichert, J. M., Moorman, T., Konopka, A. & Turco, R. F. (2002). Surface and subsurface microbial biomass, community structure and metabolic activity as a function of soil depth and season. *Applied Soil Ecology* **20**, 171-181.

Boehme, L., Langer, U. & Boehme, F. (2005). Microbial biomass, enzyme activities and microbial community structure in two European long-term field experiments. *Agriculture, Ecosystems & Environment* **In Press, Corrected Proof**.

Brosius, J., Dull, T. J., Sleeter, D. D. & Noller, H. F. (1981). Gene organization and primary structure of a ribosomal RNA operon from Escherichia coli. *Journal of Molecular Biology* **148**, 107-127.

Bruce, K. D. (1997). Analysis of mer gene subclasses within bacterial communities in soils and sediments resolved by fluorescent-PCR-restriction fragment length polymorphism profiling. *Applied & Environmental Microbiology* **63**, 4914–4919.

Christensen, S., Ronn, R., Ekelund, F. , Andersen, B., Damgaard, J., Friberg-Jensen, U., Jensen, L., Kill, H, .Larsen, B., Larsen, J. (1996). Soil respiration profiles and protozoan enumeration agree as microbial growth indicators. *Soil Biology and Biochemistry* **28**, 865-868.

Delorme, S., Philippot, L., Edel-Hermann, V., Deulvot, C., Mougel, C. & Lemanceau, P. (2003). Comparative genetic diversity of the narG, nosZ, and 16S rRNA genes in fluorescent pseudomonads. *Applied & Environmental Microbiology* **69**, 1004-1012.

Dick, R. P. (1992). A review: long-term effects of agricultural systems on soil biochemical and microbial parameters. ***Agriculture, Ecosystems & Environment*** **40**, 25-36.

Duah-Yentumi, S., Ronn, R. & Christensen, S. (1998). Nutrients limiting microbial growth in a tropical forest soil of Ghana under different management. *Applied Soil Ecology* **8**, 19-24.

Dunbar, J., Ticknor, L. O. & Kuske, C. R. (2001). Phylogenetic specificity and reproducibility and new method for analysis of terminal restriction fragment profiles of 16S rRNA genes from bacterial communities. *Applied & Environmental Microbiology* **67**, 190-197.

Emmerling, C., Udelhoven, T. & Schroeder, D. (2001). Response of soil microbial biomass and activity to agricultural de-intensification over a 10 year period. *Soil Biology and Biochemistry* **33**, 2105-2114.

Fao (1976). FAO- Unesco Soil Map of the World Legend. Unesco, Paris 1.

Fedi, S., Tremaroli, V., Scala, D., Perez-Jimenez, J. R., Fava, F., Young, L. & Zannoni, D. (2006). T-RFLP analysis of bacterial communities in cyclodextrin-amended bioreactors developed for biodegradation of polychlorinated biphenyls. *Research in Microbiology* **In Press, Corrected Proof**.

Fierer, N., Schimel, J. P. & Holden, P. A. (2003). Variations in microbial community composition through two soil depth profiles. *Soil Biology and Biochemistry* **35**, 167-176.

Fredslund, L., Ekelund, F., Jacobsen, C. S. & Johnsen, K. (2001). Development and application of a most-probable-number-PCR assay to quantify flagellate populations in soil samples. *Applied & Environmental Microbiology* **67**, 1613-1618.

Galantini, J. & Rosell, R. (2006). Long-term fertilization effects on soil organic matter quality and dynamics under different production systems in semiarid Pampean soils. *Soil and Tillage Research* **In Press, Corrected Proof**.

Galloway, J. N., Dentener, F. J., Capone, D. & other authors (2004). Nitrogen cycles: past, present, and future. *Biogeochemistry* **70**, 153-226.

Griffiths, R. I., Whiteley, A. S., O'Donnell, A. G. & Bailey, M. J. (2000). Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Applied & Environmental Microbiology* **66**, 5488-5491.

Haese, C. & Finkelstein, R. (1993). Bacterial extracellular zinc-containing metalloproteases. *Microbiological Reviews* **57**, 823-837.

Handelsman, J. (2004). Metagenomics: Application of genomics to uncultured microorganisms. ***Microbiology and Molecular Biology*** *Reviews* **68**, 669-685.

Harmsen, G. W. & Kolenbrander, G. J. (1965). Soil inorganic nitrogen. In: Harmsen, G. W. & Kolenbrander, G. J. (eds.). Soil Nitrogen American Society of Agronomy, Wisconsin, pp 43-92.

Hayano, K. (1983). Hydrolytic enzyme activities in soil materials used in nursery pot for tomato plant. ***Japanese*** *Journal of* ***Soil Science*** *and* ***Plant Nutrition*** **54**, 331-334.

Hayano, K., Takeuchi, M. & Ichishima, E. (1987). Characterization of a metalloproteinase component extracted from soil. *Biology and Fertility of Soils* **4**, 179–183.

Hayano, K., Watanabe, K. & Asakawa, S. (1995). Activity of protease extracted from rice - rhizosphere soils under double cropping of rice and wheat. ***Soil*** *Science and* ***plant Nutrition***, **41**, 597-603.

Hayano, K. (1996). Characterization and origin of protease activity in cultivated soils. *Japan Agricultural Research Quarterly* **30**, 79-84.

Head, I. M., Saunders, J. R. & Pickup, R. W. (1998). Microbial evolution, diversity, and ecology: A decade of ribosomal RNA analysis of uncultivated microorganisms. *Microbial Ecology* **35**, 1-21.

Henry, S., Baudoin, E., Lopez-Gutierrez, J. C., Martin-Laurent, F., Brauman, A. & Philippot, L. (2004). Quantification of denitrifying bacteria in soils by nirK gene targeted real-time PCR. *Journal of Microbiological Methods* **59**, 327-335.

Holt, J. A. & Mayer, R. J. (1998). Changes in microbial biomass and protease activities of soil associated with long-term sugar cane monoculture. *Biology and Fertility of Soils* **27**, 127-131.

Hoppe, H.-G., Kim, S.-J. & Gocke, K. (1988). Microbial decomposition in aquatic environments: combined process of extracellular enzyme activity and substrate uptake. *Applied & Environmental Microbiology* **54**, 784-790.

Horz, H.-P., Rotthauwe, J.-H., Lukow, T. & Liesack, W. (2000). Identification of major subgroups of ammonia-oxidizing bacteria in environmental samples by T-RFLP analysis of amoA PCR products. *Journal of Microbiological Methods* **39**, 197-204.

Jarvis, S. C., Stockdale, E. A., Shepherd, M. A. & Powlson, D. S. (1996). Nitrogen mineralization in temperate agricultural soils: processes and measurement. *Advances in Agronomy* **57**, 187-235.

Ji, R. & Brune, A. (2005). Digestion of peptidic residues in humic substances by an alkali-stable and humic-acid-tolerant proteolytic activity in the gut of soil-feeding termites. *Soil Biology and Biochemistry* **37**, 1648-1655.

Joergensen, R. G. & Castillo, X. (2001). Interrelationships between microbial and soil properties in young volcanic ash soils of Nicaragua. *Soil Biology and Biochemistry* **33**, 1581-1589.

Jones, D. L. & Kielland, K. (2002). Soil amino acid turnover dominates the nitrogen flux in permafrost-dominated taiga forest soils. *Soil Biology and Biochemistry* **34**, 209-219.

Kalisz, H. M. (1988). Microbial proteinases. *Advances in **Biochemical** Engineering/**Biotechnology*** **36**, 165.

Kamimura, Y. & Hayano, K. (2000). Properties of protease extracted from tea-field soil. *Biology and Fertility of Soils* **30**, 351-355.

Kandeler, E., Luxhoi, J., Tscherko, D. & Magid, J. (1999). Xylanase, invertase and protease at the soil-litter interface of a loamy sand. *Soil Biology and Biochemistry* **31**, 1171-1179.

Kandeler, E., Tscherko, D., Bruce, K. D., Stemmer, M., Hobbs, P. J., Bardgett, R. D. & Amelung, W. (2000). The structure and function of the soil microbial community in microhabitats of a heavy metal polluted soil. *Biology and Fertility of Soils* **32**, 390-400.

Kim, J. & Kendall, D. A. (2000). Sec-dependent protein export and the involvement of the molecular chaperone SecB. *Cell Stress & Chaperones* **5**, 267-275.

Kirk, J. L., Beaudette, L. A., Hart, M., Moutoglis, P., Klironomos, J. N., Lee, H. & Trevors, J. T. (2004). Methods of studying soil microbial diversity. *Journal of Microbiological Methods* **58**, 169-188.

Klappenbach, J. A., Dunbar, J. M. & Schmidt, T. M. (2000). rRNA Operon Copy Number Reflects Ecological Strategies of Bacteria. *Applied & Environmental Microbiology* **66**, 1328-1333.

Klappenbach, J. A., Saxman, P. R., Cole, J. R. & Schmidt, T. M. (2001). rrndb: the ribosomal RNA operon copy number database. *Nucleic Acids Research* **29**, 181-184.

Kloos, K., Husgen, U. & Bothe, H. (1998). DNA-probing for genes coding for denitrification, N2-fixation and nitrification in bacteria isolated from different soils. *Zeitschrift für **Naturforschung*** **53**, 69-81.

Kristufek, V., Elhottova, D., Chronakova, A., Dostalkova, I., Picek, T. & Kalcik, J. (2005). Growth strategy of heterotrophic bacterial population along successional sequence on spoil of brown coal colliery substrate. *Folia Microbiologica (Praha)* **50**, 427-435.

Ladd, J. N. & Butler, J. H. A. (1972). Short-term assays of soil proteolytic enzyme activities using proteins and dipeptide derivatives as substrates. *Soil Biology and Biochemistry* **4**, 19-30.

Law, B. A. (1980). Transport and utilization of proteins by bacteria. In: Payne J.W. (eds.). Microorganisms and Nitrogen Sources. John Wiley and Sons, London, pp 381-409.

Liesack, W., Janssen, P. H., Rainey, F. A., Ward-Rainey, N. L. (1997). Microbial diversity in soil: the need for a combined approach using molecular and cultivation techniques. In: van Elsas, J. D., Trevors J. T., and Wellington E. M. H. (eds.). Modern Soil Ecology. Marcel Dekker, New York, pp 375-439.

Lipson, D. & Näsholm, T. (2001). The unexpected versatility of plants: organic nitrogen use and availability in terrestrial ecosystems. *Oecologia* **128**, 305-316.

Loll, M. J. & Bollag, J.-M. (1983). Protein transformation in soil. *Advances in Agronomy* **36**, 351-382.

Lopez-Hernandez, D., Nino, M., Nannipieri, P. & Fardeau, J. C. (1989). Phosphatase activity in Nasutitermes ephratae termite nests. *Biology and Fertility of Soils* **7**, 134-137.

Lorenz, K. & Kandeler, E. (2005). Biochemical characterization of urban soil profiles from Stuttgart, Germany. *Soil Biology and Biochemistry* **37**, 1373-1385.

Lueders, T. & Friedrich, M. (2000). Archaeal population dynamics during sequential reduction processes in rice field soil. *Applied & Environmental Microbiology* **66**, 2732-2742.

Mantynen, V., Niemela, S., Kaijalainen, S., Pirhonen, T. & Lindstrom, K. (1997). MPN-PCR-quantification method for staphylococcal enterotoxin c1 gene from fresh cheese. *International Journal of Food Microbiology* **36**, 135-143.

Marshman, N. A. & Marshall, K. C. (1981). Bacterial growth on proteins in the presence of clay minerals. *Soil Biology and Biochemistry* **13**, 127-134.

Marx, M.-C., Kandeler, E., Wood, M., Wermbter, N. & Jarvis, S. C. (2005). Exploring the enzymatic landscape: distribution and kinetics of hydrolytic enzymes in soil particle-size fractions. *Soil Biology and Biochemistry* **37**, 35-48.

Mumy, K. L. & Findlay, R. H. (2004). Convenient determination of DNA extraction efficiency using an external DNA recovery standard and quantitative-competitive PCR. *Journal of Microbiological Methods* **57**, 259-268.

Murphy, D. V., Macdonald, A. J., Stockdale, E. A. & other authors (2000). Soluble organic nitrogen in agricultural soils. *Biology and Fertility of Soils* **30**, 374-387.
Neher, D. A., Barbercheck, M. E., El-Allaf, S. M. & Anas, O. (2003). Effects of disturbance and ecosystem on decomposition. *Applied Soil Ecology* **23**, 165-179.

Newell, S. Y., Arsuffi, T. L. & Palm, L. A. (1996). Misting and nitrogen fertilization of shoots of a saltmarsh grass: effects upon fungal decay of leaf blades. *Oeacologia* **108**, 495-502.

Pace, N. R., Stahl, D. A., Lane, D. J. & Olsen, G. J. (1986). The analysis of natural microbial populations by ribosomal RNA sequences. *Advances in Microbial Ecology* **9**, 1-55.

Paul, E. A. & Juma, N. G. (1981). Mineralization and immobilization of nitrogen by microorganisms.In: Clark, F. E. and Rosswall, T. (eds.). Terrestrial nitrogen cycles. Ecology Bulletin, Stockholm, pp 179-195.

Paul, E. A. & Clark, F. E. (1996). Ammonification and nitrification. In: Paul, E. A. & Clark, F. E. (eds.) Soil Microbiology and Biochemistry. Academic Press, San Diego, pp. 182-183

Qiu, X., Wu, L., Huang, H., McDonel, P. E., Palumbo, A. V., Tiedje, J. M. & Zhou, J. (2001). Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. *Applied & Environmental Microbiology* **67**, 880-887.

Qiuhong, N., Xiaowei, H., Baoyu, T., Jinkui, Y., Jiang, L., Lin, Z. & Keqin, Z. (2006). *Bacillus* sp. B16 kills nematodes with a serine protease identified as a pathogenic factor. *Applied Microbiology and Biotechnology* **69**, 722-730.

Rahman, R. N. Z. A., Geok, L. P., Basri, M. & Salleh, A. B. (2005). Physical factors affecting the production of organic solvent-tolerant protease by Pseudomonas aeruginosa strain K. *Bioresource Technology* **96**, 429-436.

Rao, M. B., Tanksale, A. M., Ghatge, M. S. & Deshpande, V. V. (1998). Molecular and biotechnological aspects of microbial proteases. *Microbiology and Molecular Biology Reviews* **62**, 597-635.

Rich, J. J. & Myrold, D. D. (2004). Community composition and activities of denitrifying bacteria from adjacent agricultural soil, riparian soil, and creek sediment in Oregon, USA. *Soil Biology and Biochemistry* **36**, 1431-1441.

Rousselon, N., Delgenes, J.-P. & Godon, J.-J. (2004). A new real time PCR (TaqMan(R) PCR) system for detection of the 16S rDNA gene associated with fecal bacteria. *Journal of Microbiological Methods* **59**, 15-22.

Rowell, M. J., Ladd, J. N. & Paul, E. A. (1973). Enzymatically-active complexes of proteases and humic acid analogues. *Soil Biology and Biochemistry* **5**, 699-703.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406-425.

Sardans, J. & Penuelas, J. (2005). Drought decreases soil enzyme activity in a Mediterranean Quercus ilex L. forest. *Soil Biology and Biochemistry* **37**, 455-461.

Schlichting, E., Blume, H.-P. & Stahr, K. E. (1995). Bodenkundliches Praktikum. Parey Verlag, Berlin, pp. 295

Schloter, M., Bach, H.-J., Metz, S., Sehy, U. & Munch, J. C. (2003). Influence of precision farming on the microbial community structure and functions in nitrogen turnover. *Agriculture, Ecosystems & Environment* **98**, 295-304.

Schmidt, B., Woodhouse, L., Adams, R., Ward, T., Mainzer, S. & Lad, P. (1995). Alkalophilic Bacillus sp. strain LG12 has a series of serine protease genes. *Applied & Environmental Microbiology* **61**, 4490-4493.

Schroeder, P., Huber, B., Olazabal, U., Kaemmerer, A. & Munch, J. C. (2002). Land use and sustainability: FAM research network on agroecosystems. *Geoderma* **105**, 155- 166.

Schulten, H. R. & Schnitzer, M. (1998). The chemistry of soil organic nitogen: a review. *Biology and Fertility of Soils* **26**, 1-15.

Schwieger, F. & Tebbe, C. C. (1998). A new approach to utilize PCR-single-strand-conformation polymorphism for 16S rRNA gene-based microbial community analysis. *Applied & Environmental Microbiology* **64**, 4870-4876.

Scott Bechtold, J. & Naiman, R. J. (2006). Soil texture and nitrogen mineralization potential across a riparian toposequence in a semi-arid savanna. *Soil Biology and Biochemistry In Press, Corrected Proof.*

Sessitsch, A., Weilharter, A., Gerzabek, M. H., Kirchmann, H. & Kandeler, E. (2001). Microbial population structures in soil particle size fractions of a long-term fertilizer field experiment. *Applied & Environmental Microbiology* **67**, 4215-4224.

Sharma, S. (2003). Structural and functional characterization of bacterial diversity in the rhizospheres of three grain legumes. In: Dissertation zur Erlangung des Doktorgrades der Fakultät für Biologie der Ludwig-Maximilians-Universität Muenchen, Muenchen, pp 18-19.

Sharma, S., Aneja, M. K., Mayer, J., Schloter, M. & Munch, J. C. (2004). RNA fingerprinting of microbial community in the rhizosphere soil of grain legumes. *FEMS Microbiology Letters* **240**, 181-186.

Sharma, S., Aneja, M. K., Mayer, J., Munch, J. C. & Schloter, M. (2005). Diversity of transcripts of nitrite reductase genes (nirK and nirS) in rhizospheres of grain legumes. *Applied & Environmental Microbiology* **71**, 2001-2007.

Sims, G. K. & Wander, M. M. (2002). Proteolytic activity under nitrogen or sulfur limitation. *Applied Soil Ecology* **19**, 217-221.

Skujins, J. (1976). Extracellular enzymes in soil. *CRC Critical Review Microbiology* **4**, 383-421.

Smit, E., Leeflang, P., Gommans, S., van den Broek, J., van Mil, S. & Wernars, K. (2001). Diversity and seasonal fluctuations of the dominant members of the bacterial soil community in a wheat field as determined by cultivation and molecular methods. *Applied & Environmental Microbiology* **67**, 2284-2291.

Sorensen, J. (1997). The rhizosphere as a habitat for soil microorganisms. In: van Elsas, J. D., Wellington, E. M. H. and Trevors, J. T. (eds.). Modern Soil Microbiology, Marcel Dekker, New York, pp 21-46.

Stevenson, F. J. (1982). Organic forms of soil nitrogen. In: Stevenson F.J. (ed.). Nitrogen in agricultural soils. American Society of Agronomy, Madison, pp 67-122.

Takagi, M., Imanaka, T. & Aiba, S. (1985). Nucleotide Sequence and Promoter Region for the Neutral Protease Gene from *Bacillus stearothermophillus*. *Journal of Bacteriology* **163**, 824-831.

Taylor, J. P., Wilson, B., Mills, M. S. & Burns, R. G. (2002). Comparison of microbial numbers and enzymatic activities in surface soils and subsoils using various techniques. *Soil Biology and Biochemistry* **34**, 387-401.

Team, R. D. C. (2005). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, **URL http://www.R-project.org.**, ISBN 3-900051-900007-900050.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680.

Van Gestel, M., Ladd, J. N. & Amato, M. (1992). Microbial biomass responses to seasonal change and imposed drying regimes at increasing depths of undisturbed topsoil profiles. *Soil Biology and Biochemistry* **24**, 103-111.

Vitousek, P. M. & Howarth, R. (1991). Nitrogen limitation on land and in the sea: how can it occur? *Biogeochemistry* **13**, 87-115.

Wandersman, C. (1989). Secretion, processing and activation of bacterial extracellular proteases. *Molecular Microbiology* **3**, 1825-1831.

Wardle, D. A. (1998). Controls of temporal variability of the soil microbial biomass: A global-scale synthesis. *Soil Biology and Biochemistry* **30**, 1627-1637.

Watanabe, K. & Hayano, K. (1994). Estimate of the source of soil protease in upland fields. *Biology and Fertility of Soils* **18**, 341-346.

Watanabe, K. & Hayano, K. (1995). Seasonal variation of soil protease activities and their relation to proteolytic bacteria and Bacillus spp in paddy field soil. *Soil Biology and Biochemistry* **27**, 197-203.

Watanabe, K., Sakai, J. & Hayano, K. (2003). Bacterial extracellular protease activities in field soils under different fertilizer managements. *Canadian Journal of Microbiology* **49**, 305-312.

Widmer, F., Shaffer, B. T., Porteous, L. A. & Seidler, R. J. (1999). Analysis of nifH gene pool complexity in soil and litter at a douglas fir forest site in the Oregon cascade mountain range. *Applied & Environmental Microbiology* **65**, 374-380.

Wolsing, M. & Prieme, A. (2004). Observation of high seasonal variation in community structure of denitrifying bacteria in arable soil receiving artificial fertilizer and cattle manure by determining T-RFLP of nir gene fragments. *FEMS Microbiology Ecology* **48**, 261-271.

Yallop, C. A., Edwards, C. & Williams, S. T. (1997). Isolation and growth physiology of novel thermoactinomycetes. *Journal of Applied Microbiology* **83**, 685-692.

Zeller, V., Bardgett, R. D. & Tappeiner, U. (2001). Site and management effects on soil microbial properties of subalpine meadows: a study of land abandonment along a north-south gradient in the European Alps. *Soil Biology and Biochemistry* **33**, 639-649.

Zhou, J., Xia, B., Huang, H., Palumbo, A. V. & Tiedje, J. M. (2004). Microbial diversity and heterogeneity in sandy subsurface soils. *Applied & Environmental Microbiology* **70**, 1723-1734.

7. APPENDIX

7.1 Figure legends

Figure 1 Nitrogen cycle (modified from K. Kloos, not published data) 3

Figure 2 Heterogeneity and the main soil types of the field site (A 15) investigated in the present study 12

Figure 3 Diagram of the Real-Time PCR amplification plot (http:// dna9.intmed .iouwa.edu/realtime.htm) 20

Figure 4 Principle of T-RFLP approach (http:// rdp8.cme.msu.edu/html/t-rflp/ 02.html) 23

Figure 5 WEOC measurements investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard deviation of means. 35

Figure 6 WEON measurements investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard deviation of means. 36

Figure 7 DNA concentration measurements investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. 37

Figure 8 Obtained *nprA* PCR products in soil samples with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) electrophoretically separated on 1.5 % agarose gel. As an exemplar results from October were shown. 38

Figure 9 Obtained *sub* PCR products in soil samples with four different soil types (GC, LC, CP, and GK) and three soil depths (D1D3) electrophoretically separated on 1.5 % agarose gel. As an exemplar results from October were shown. 38

Figure 10 Standard curves for quantification of soil a) *npr*, b) *sub* and c) 16S rRNA. Black dots represent standard DNA with known copies number and rot represent soil DNA with unknown gene copies number. 40

Figure 11 Number of total (A-C), *npr* (D-F) and *sub* (G-I) coding proteolytic bacteria investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard errors of means. 43

Figure 12 Potential proteolytic activity investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard errors of means. 44

Figure 13 Relationship between *sub* copy number and potential proteolytic activity at site/depth LC D3 (A) and site/depth CP D3 (C). Relationship between *npr* copy number and potential proteolytic activity at site/depth LC D3 (B) and site/depth CP D3 (D) in April, July and October 2003. (SQRT=square root transformed data) 46

Figure 14 Electropherogram of the *sub* T-RFLP from (A) *Bacillus amyloliquefaciens* DSM 7, (B) *Bacillus licheniformis* DSM 13, (C) *Bacillus subtilis* DSM 10 47

Figure 15 Electropherogram of the *npr* T-RFLP from (A) *Bacillus stearothermophilus* DSM 22, (B) *Bacillus thuringiensis* DSM 2046, (C) *Bacillus sp*. DSM 405 (D) *Bacillus cereus* DSM 3101 (E) *Bacillus megaterium* DSM 90 48

Figure 16 Number of *npr* T-RFs investigated in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard deviation of means. 50

Figure 17 Length (bp) and relative abundance (%) of dominated *npr* T-RFs in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. 50

Figure 18 PCA ordering of data (axis 1 und 2) generated from *npr* T-RFLP fragments amplified from soil samples collected at field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. A symbol represents median of three measurements. Circles present GC, triangles LC, rectangles CP and diamonds GK. Gray represents April, black July and white October. The numbers are as followed: 1 for D1, 2 for D2 and 3 for D3. 51

Figure 19 Number of *sub* T-RFs investigated in soil samples from field site with four different soil types (GC, LC, CP and GK) and three soil depths (D1-D3) in April, July and October 2003. Error bars represent standard deviation of means. 52

Figure 20 Length (bp) and relative abundance (%) of dominated *sub* T-RFs in soil samples from field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. 53

Figure 21 PCA ordering of data (axis 1 und 2) generated from *sub* T-RFLP fragments amplified from soil samples collected at field site with four different soil types (GC, LC, CP, and GK) and three soil depths (D1-D3) in April, July and October 2003. A symbol represents median of

three measurements. Circles present GC, triangles LC, rectangles CP and diamonds GK. Gray represents April, black July and white October. The numbers are as followed: 1 for D1, 2 for D2 and 3 for D3. 54

Figure 22 Neighbor-joining analysis of *npr* sequences based on 120 amino acids cloned from the soil samples. Clones obtained from this investigation are shown with soil type labeling and the number of the clone under respective soil type library. Bootstrap values for a total of 1000 replicates are given, values below 500 are omitted. A single sequence was taken to represent all sequences in cluster consisting of > 98% identical sequences: [a] GK-59, CP-3, 7, 34, 54, f; [b] GC-29; [c] GC-35, CP-2, 56; [d] CP- 6, 9; [e] GC-8; [f] GC-26; [g] GC-57, GK-49; [h] LC-30; [i] LC-3; [j] LC-16, 28, 36; [k] GK-58, CP-39; [l] GK-16. 57

Figure 23 Rarefaction analysis of the 120 *npr* clones in soil samples from field site with four different soil types (GC, LC, CP, and GK) in July 2003. Calculation was preformed at the species level (> 98% sequence similarity). 58

Figure 24 Rarefaction analyses of the *npr* clones in soil samples from field site with four different soil types (GC, LC, CP, and GK) in July 2003. Calculations were preformed at the species level (> 98% sequence similarity). 59

7.2 Table legends

Table 1 Management of the field A15 in Scheyern in 2003 13

Table 2 Bacteria used for microbiological analysis 15

Table 3 Primers and cycling conditions used in this study 17

Table 4 Characteristics of oligonucleotides used as primers 17

Table 5 Amounts of different components used in the amplification reaction 18

Table 6 16S rRNA specific TaqMan- probe 22

Table 7 Origin of *sub* specific sequences utilized for NEBcutter analysis and expected fragment size after restriction endonuclease digestion of 319 bp long *sub*- amplicon 24

Table 8 Origin of *npr* specific sequences utilized for NEBcutter analysis and expected fragment size after restriction endonuclease digestion of 233 bp long *npr*- amplicon 25

Table 9 Program cycle conditions of T-RFLP analysis on CEQ^{TM} 2000 XL DNA Analysis System 27

Table 10 Some physicochemical properties of the soil horizons sampled at the investigated field 34

Table 11 DNA extraction efficiency of soil samples 41

Table 12 Pearson correlation matrix between microbiological and chemical properties in soil samples from field site with four different sites and three different depths in April, July and October 2003. 45

Table 13 Expected and observed lengths of *sub* T-RFs for the tree organisms possessing gene coding for Sub 48

Table 14 Expected and observed lengths of *npr* T-RFs for the five organisms possessing gene coding for Npr 49

8. Acknowledgements

My cordial thanks to:

Prof. Dr. J. C. Munch, for giving me the opportunity to work in this project, especially for his interest in the progress of the research and for the constructive discussions for better understanding of the subject.

Dr. Michael Schloter, for providing comfortable work environment and for his suggestions during the course of the work. Thank you for being there to discuss whenever needed.

Dr. Marion Engel, for careful supervision, new and constructive ideas that enrich my thesis and moreover for being there as a scientist and a friend.

Prof. Dr. Scherer and Prof. Dr. I. Kögel-Knabner for their willingness to review this work.

Dr. Karin Pritsch, for being there always at the right moment to encourage me to overcome difficult situations. I am very grateful for having the opportunity to take a little piece of your knowledge. I would like to thank you for everything but mostly to be there as a friend.

Dr. Alexandra Hagn and Dr. Viviane Radl, for being there to solve my scientific problems but also to be there when I needed laugh, break and friendly conversation. I would like to thank you for standing open door of your 5-m^2 office where lots of scientific and non-scientific discussions took place.

Dr. Shilpi Sharma, for constructive scientific discussions and optimism.

Dr. Andreas Gattinger, for scientific discussions and his help in the preparations of manuscripts.

Ms. C. Galonska, for her expert assistance with laboratory techniques.

Great thanks for the AG Schloter for friendly working environment, constructive scientific discussions and ever helping attitude.

My husband and best friend Ivor Fuka for his love, patience and support.

My parents for the unconditional support they gave me all those years.