UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN
Scienze Agrarie e Ambientali

CICLO XXVIII

COORDINATORE Prof. Stefano Mancuso

Microbial protease encoding genes in soil: diversity, abundance and enzymatic activity

Settore Scientifico Disciplinare AGR/13

Dottorando
Dott. Divyashri Baraniya

Tutore
Prof. Paolo Nannipieri

and Prof. Giacomo Pietramellara

Coordinatore
Prof. Stefano Mancuso

Anni 2012/2015

Dedicated to

My Grandfather

Who has always been a source of constant inspiration and encouragement

"Have no fear of perfection; you'll never reach it."

"Nothing in life is to be feared; it is only to be understood."

Marie Skłodowska Curie

Abstract

Enzymatic activities in soil are an important component of soil health and biodiversity. Proteases catalyse the breakdown of proteins and peptidases resulting in release of amino acids, which can be used by plants, microbes and other organisms for meeting their N requirement. Microbes are the largest source of proteases in soil. We hypothesized that the protease activity and proteolytic communities in soil gets affected by environmental and biotic factors. The objective of this research was to identify potential factors that govern the abundance, diversity and expression of genes encoding proteases. We studied the response of protease encoding genes and proteolytic community structure of soil under different conditions.

Measurement of enzyme activity was carried out by colorimetric determination of enzyme-substrate complex after fixed incubation. To investigate diversity and structure of proteolytic communities, PCR-DGGE as well as high throughput Illumina sequencing of amplicons were used. Abundance of genes was studied using qPCR. Expression of protease encoding genes was studied using Illumina sequencing of soil metatranscriptome. Three different approaches have been used for our studies:

1. Using a rhizobox approach effect of root exudates from plants differing in Nitrogen utilizing efficiencies (NUE), on protease enzyme activity, molecular diversity and abundance of proteolytic genes was investigated. We observed a higher molecular diversity, abundance and higher enzyme activity associated with higher NUE cultivar. Furthermore, effect of root exudates can't be ignored and rhizosphere soil exhibited significantly higher results. Illumina sequencing results show that the OTUs from *Bacillus spp.* was dominant among *npr*

protease gene, wherease *Pseudomonas spp.* was major source of *apr* protease gene. We also noticed that most proteolytic bacterias were also PGPRs.

2. Effect of elevated atmospheric CO_2 was studied on rhizosphere and bulk soil from a Free Air Carbon dioxide Enrichment (FACE) field that was exposed to 550 ppm of CO_2. We observed a reduction in abundance of apr protease genes. This decrease has been observed in both rhizosphere and bulk soil. We concluded that in eCO2 conditions, the action on soil microbes is not limited via root exudates but also mediated through fixation of CO2 directly by mircobes under high partial pressure.

3. Changes in expression of proteases during day and night has been studied in a green-house experiment where Barley rhizosphere soil was subjected to metatranscriptome analysis. Samples were collected during dark and light exposed period and mRNA was isolated which was later sequenced on an Illumina sequencer. Bioinformatics analysis is ongoing, but preliminary results suggest a significant difference in amino acid metabolism pathway and shifts in some microbial orders during the light exposed period.

Results from above three approaches in a nutshell led to conclusion that protease enzyme activity, proteolytic gene diversity and abundance are not only a direct outcome of microbial activity but root exudates from plants have a strong influence on altering microbial proteolytic community structure, and their diversity and abundance. Any environmental or biotic factor affecting plants also affect the proteolytic gene profile and its expression in soil. Apart from root exudates eCO_2 was also observed to control abundance of microbial population and diversity, and a reduction in genes responsible for proteolytic activity in soil treated with eCO_2 was observed. Day and night cycles of plants also affect the microbial community structure and functions related to amino acid metabolism and proteolysis and

an increased protease expression was observed in Barley Rhizosphere soil in night time than in day.

Contents

List of abbreviations

AMC	7-amino 4-methyl Coumarin
NH_4^+-N	Ammonium nitrogen
ANOVA	Analysis of Variance
BLAST	Basic Local Alignment Search Tool
BAA	benzoyl-L-argininamide
C	Carbon
DGGE	Denaturing gradient gel electrophoresis
EC	Enzyme class
FACE	Free Air Carbon dioxide Enrichment
KEGG	Kyoto Encyclopedia of Genes and Genomes
mRNA	messenger RNA
MPN	most probable number
MTB	*Mycobacterium tuberculosis*
ZPL	N-benzyloxy-carbonyl-L-phenylalanyl L-leucine
N	Nitrogen
NC-IUBMB	Nomenclature committee of International Union of Biochemistry and Molecular Biology
NTM	*Non-tuberculous Mycobacteria*
PCR	Polymerase chain reaction

PCA	Principal Component Analysis
PDB	Protein data bank
Rr	Range weighted richness
ROS	Reactive Oxygen Species
SOM	Secondary Organic Matter
TRFLP	Terminal restriction fragment length polymerism
UPGAMA	Unweighted Pair-Group Method With Arithmetic Averages

1. Introduction

Nitrogen is an important element that is required by plants, animals and microbes for their growth and maintaining cell structure. In soil Proteins, Chitin and Urea are major forms of organic N. In soil largest fraction of organic N is constituted by proteins. Proteases and peptidases are enzymes that carry out hydrolysis of proteins and peptides, respectively by cleaving the peptide bond, thus releasing N as amino acids, which may be mineralized to inorganic N to be used by plants (Nannipieri and Paul, 2009). Proteases and peptidases control many important processes inside and outside cells, and have diverse biological functions apart from controlling dynamics of protein and N turnover. Protease activity is one of the many enzyme activities considered as an indicator of soil health. Activity of any enzyme is an outcome of the response of cellular machinery mediated via genes. Thus protease activity is an outcome of behaviour of proteolytic genes. Despite such an important role in soil functioning, there is little understanding on effects of different biological and environmental effects on protease activity and on diversity, abundance and expression of genes encoding proteases. Understanding of various aspects and factors affecting soil proteases and protease encoding genes can help in understanding soil functions.

1.1 Classification

The classification of proteases, based on their catalytic site and action mechanism, gives 4 major groups: Serine proteases, thiol proteases, acid proteases and metal proteases

(Hartley, 1960). But with a better identification of the catalytic mechanisms and specificity, proteases and peptidases have at present different classifications based on:

(a) Active site and catalytic mechanism: Based on this, peptidases have been assigned 9 classes: serine, cysteine, threonine, aspartic, glutamic, asparagine or metallo-catalytic type and mixed catalytic type. (MEROPS database: https://merops.sanger.ac.uk/about/classification.shtml).

(b) Reaction catalysed: endopeptidases, omega-peptidases, carboxypeptidases, dipeptidyl-peptidases, tripeptidyl-peptidases, peptidyl-dipeptidases and dipeptidases (Rawlings et al. 2012).

(c) Molecular structure and homology: This classification is based more on mechanism of action, specificity and physiological action (Rawlings and Barrett, 1993). According to this classification, proteases and peptidases are grouped in families or in a group of related families called clan.

(d) Optimum pH requirement: based on their optimal pH proteases are classified as: acid proteases, alkaline proteases and neutral proteases.

The Nomenclature committee of International Union of Biochemistry and Molecular Biology (NC-IUBMB), has placed peptidases in Enzyme Class (EC): EC 3.4; Class 3 include hydrolases (http://www.chem.qmul.ac.uk/iubmb/enzyme/EC34/). Based on site of action Proteases are broadly subdivided in two sub-subclasses: exopeptidases (EC 3.4.11-19) and endopeptidases (EC 3.4.21-24 and EC 3.4.99).

Exopeptidases targets terminal peptide bonds and cleave di and tri-peptides into their constituent amino acids, whereas endopeptidases attack peptide bonds of non-terminal amino acids. Exodpeptidases and endopeptidases are further divided in many sub-subclasses based on their action mechanism and active site as (Table1).

1.2 General mechanism

Because of involvement of proteases in intracellular and extracellular, these enzymes have been extensively studied. Many attempts have been made to understand the action mechanism of specific types of protease. Mechanism of action depends on the type of protease and its catalytic site (Williams, 1969, Drenth, 1980, Antonov et al. 1981, Nessi et al., 1998, Polgár 2005). In general hydrolysis of the peptide bond involves the formation of a tetrahedral intermediate and is an addition-elimination reaction, mediated by a nucleophilic attack on peptide bond. Serine, cysteine and threonine proteases use a nucleophile from a catalytic triad to perform a nucleophilic attack and this involves formation of an acyl – covalent intermediate. On the other hand aspartic, glutamic and metallo-proteases activate water molecule to carry out a nucleophilic attack on the peptide bond (Polgár, 1989).

Proteases and peptidase
(EC 3.4)

Exopeptidases (EC 3.4.11-19)		Endopeptidases (EC 3.4.21-25 and EC 3.4.99)	
Aminopeptidases (3.4.11)	Release single amino acid by acting at free N– terminus of a polypeptide chain	Serine endopeptidases (3.4.21)	
Dipeptidases (3.4.13)	This exodpeptidases is specific for dipeptide	Cysteine endopeptidases (3.4.22)	
Dipeptidyl-peptidases and tripeptidyl-peptidases (3.4.14)	Release dipeptide or tripeptide by acting at free N terminus of a polypeptide chain	Aspartic endopeptidases (3.4.23)	
Peptidyl-dipeptidases (3.4.15)	Release a dipeptide by acting at free C-terminus of a polypeptide	Metalloendopeptidases (3.4.24)	
Carboxypeptidases (3.4.16-18)	Release a dipeptide by acting at free C-terminus of a polypeptide. Based on the type of catalytic mechanism, these are of 3 types	Threonine endopeptidases (3.4.25)	
Serine-type Carboxypeptidases (3.4.16)		Endopeptidases of unknown catalytic mechanism (3.4.99)	These endopeptidases can't be assigned in any of the above sub class
Metallocarboxypeptidases (3.4.17)			
Cysteine-type carboxypeptidases (3.4.18)			
Omega peptidases (3.4.19)	Act on terminal residues that are substituted, cyclized or linked by isopeptide bonds		

Table1. Classification of proteases according to NC-IUBMB

1.3 Databases:

MEROPS, is a well-known database dedicated to peptidases, proteases, proteinases and protease inhibitors and the classification is based on the evolutionary relatedness (Rawlings and Barrett, 1993). Since its introduction (Rawlings and Barrett, 1999), this database has been updated from time to time (Rawlings et al 2004a, 2004b, 2006, 2008, 2010, 2012 and 2014, Barrett et al. 2001). 50 clans of peptidases are recognized by the recent version 9.9 of the MEROPS database. Peptidase search is possible by its peptide name (Partial or full), or by a known accession number belonging to a database like EMBL/GenBank, RCSC-PDB, SwissProt, TrEMBL, and PIR.

ProLysED is a metaserver integrated database for bacterial protease systems and its dataset also includes regulatory and inhibitory proteins acting on proteases (Firdaus et al. 2005).

MycoProtease-DB, is a database strictly for proteases belonging to selected 12 known tuberculosis causing bacteria strains, whose complete genome has been sequenced (Jena et al. 2012). Of these 8 strains belongs to *Mycobacterium tuberculosis* (MTB) and 4 strains belongs to Nontuberculous Mycobacteria (NTM).

PMAP-CutDB maintains a data of curated proteolytic events, thus not only it contains information about proteases but also other associated information with proteolysis like pathway, substrates and structural profile of proteases.

Apart from above mentioned databases Swissprot, NCBI, EMBL, RCSC-PDB also maintain a collection of sequences belonging to proteases, but none of these is protease specific database like MEROPS.

1.4 Roles of proteases in soil

Both classical and modern analytical techniques have indicated that the most prevalent form of organic N in soil is protein (Nannipieri and Paul, 2009). Initial breakdown of proteins from the soil organic matter is virtually mediated by soil proteases. Break down of proteins to constituent amino acids by proteases bring about N mineralization in soil (Ladd and Jackson, 1982), thus they play an important role in making N available in forms that can be readily used by plants, microbes and other soil fauna for their growth and nutrition. Extracellular proteases regulate the releases of NH4+-N thus regulating the terrestrial N cycle (Sardans et al 2008).

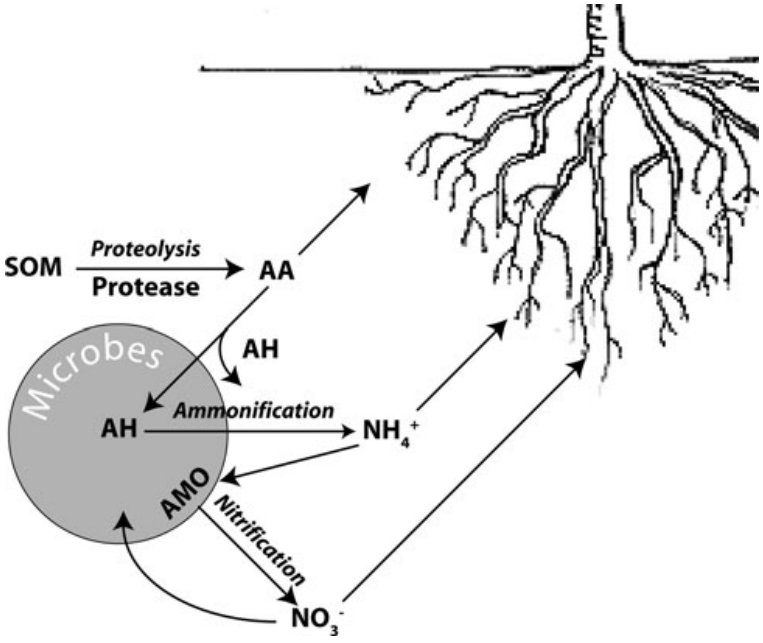


Figure 1 N mineralization brought about by proteases and incorporation of N in the terrestrial N cycle (Hofmockel et al. 2010)

Soil protease activity not only represents the proteolytic potential of soil but may also have a role in ecology of soil microbes (Burns 1982). These enzymes also play an important role in maintaining soil health (Das and Varma, 2011). Application of exogenous protease in agricultural fields showed that extracellular proteases brings down N deficiency and thus help in improving soil fertility (Han and He, 2010).

1.5 Sources

In soil both higher plants and microbes are sources of extracellular proteases. Plants secrete large amounts of proteases during their early growth stages to get sufficient N from soil. Seedlings of some plant species have shown to exudate proteases (Godlewski and Adamczyk 2007, Adamczyk et al 2007). Not only seedlings but plant themselves too are capable of exudating proteases for their growth and can show independent protease exudation, without any involvement of microbes and other soil fauna (Paungfoo-Lonhienne et al., 2008). Biochemical properties of proteases from root exudates are different for different plant species, but have a common role that is making N available for meeting the N nutrition demand (Adamczyk et al 2010).

Despite plants are higher life forms, microbes leave plants far behind in production of extracellular proteases and are major source of proteases in soil (Pansombat et al. 1997). Among microbes certain fungi especially Aspergillus spp. cannot be ignored as contributor to soil extracellular protease pool (Oseni, 2011, Choudhary and Jain, 2012, Kamath et al. 2010). Though fungal production of protease cannot be ignored, Bacteria are more

voracious producers of proteases (Sharma et al. 2015), and these are actually bacteria among microbes that outnumber any other form and are major sources of extracellular proteases in soil (Watanbe and Hayano, 1993 and 1994). In Andosol fields members of Bacillus spp. are evidenced to be a dominant source of protease (Watanabe and Hayano, 1994). Bach and Munch (2007) revealed that Pseudomonas, Bacillus and Flavobacterium are most important proteolytic species in soil. In soil proteases are also associated with organic colloids, humic acid and often also found in immobilized forms bound to clay and humic acids (Burns 1982). Often protease-humic acid complexes also show significant activity in soil (Rowell et al. 1973), but binding of enzymes to clay particles results in a decreased bacterial degradation or reduced activity (Marshman and Marshall, 1981).

1.6 Determination of microbial proteases

Most common methods of proteolytic activity determination in soil involves colorimentric or fluormetric assays, as these assays are quick and senstivite. Enzyme activity is detected by either: (i) measuring the decrease in initial substrate or (ii) measuring the increase of amino acids or peptides released during the incubation period. In colorimetric activity assays the change in the absorbance of the soil-suspenion after incubation with suitable substrate is measured. Some of the commonly used substrates for soil proteases and peptidases are N-benzoyl-Largininamide (BAA) which is specific for trypsin, N-benzyloxy-carbonyl-L-phenylalanyl L-leucine (ZPL) specific for carboxypeptidases, and casein which is essentially non-specific (Ladd and Butler 1972, Bonmati et al. 1998). After incubation with suitable substrate for a specific time a coloring reagent like Folin's reagent is added and absorbance is measured on a spectrophotometer at a specified wavelength. Another method for

measuring soil protease activity involve use of fluorogenic substrate derived from 7-amino 4-methyl Coumarin (AMC). In this approach protease activity is measured as the hydrolysis rate of L-Leucine 7-amido-4-methyl coumarin hydrochloride (Hendel and Marxsen, 2005, Brankatschk et al, 2011). Here the fluorescence is measured on a fluorometer. Both colorimetric and fluorimetric method can be performed using a cuvette or for many samples microtiter well plates can be used. To identify nature of soil protease isolelectric focusing is also used (Hayano et al, 1987).

1.7 Isolation of soil proteolytic bacteria and genes

In some of the previous studies culture dependent methods were used to isolate proteolytic bacteria. These methods include use of selective media that often involve gelatin incorporation in medium (Watanabe and Hayano, 1994b). In some advanced approaches to estimate number of proteolytic organisms in soil most probable number (MPN) and plate counting were also used (Watanabe and Hayanao, 1995, Bach and Munch, 2000). Culture based methods, though need less resources, are tedious and not always reliable as very few of total microbes are culturable.

With advances in molecular techniques, PCR primers and probes were developed for identification of some key extracellular soil proteases of bacterial origin (Bach et al. 2001). These primers allowed to study with better accuracy the diversity and distribution of proteolytic genes in agricultural soil. These primers were used to study the abundance of the proteolytic genes in soils from different sources using qpCR (Bach et al. 2002, Rasche et al. 2014). 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).

Diversity of *apr* and npr protease genes were also studied using a DGGE approach (Sakurai et al. 2007) using the primers developed by Bach et al. 2001. Fuka et al. 2009 used Terminal restriction fragment length polymerism (TRFLP), to study diversity of npr protease genes in an agricultural soil. TRFLP 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. Molecular techniques are more rapid, reliable and sensitive than culture based techniques to quantify and detect proteolytic communities. To gain detailed information about the composition of the proteolytic microbes, npr gene PCR products were also cloned and sequenced. Tsuboi et al. 2014, also sequenced cloned npr genes. However, not necessarily the fragments used for quantitative PCR (qPCR) or fingerprinting are well suited for sequencing of cloned fragments and their phylogenetic analysis as the length of the generated sequences is too short, also not percentage of successful inserts is also doubtful thus cloned sequences doesn't represent all amplified sequences. Furthermore this approach is not suitable for high throughput analysis. There is scope of using better sequencing approaches like high throughput Illumina sequencing of amplicons, study of metagenome and metatranscriptome. Metagenome represent all the DNA and gene sequences from microbial community of an environmental sample and likewise metatranscriptome represent all expressed genes from a community.

1.8 Hypotheses and Objectives

There are many environmental and biotic factors that govern the structure of microbial communities and their function in soil. The effect of these factors is poorly known. We hypothesize that the changes in soil protease activity and diversity and abundance of protease encoding genes take place due to changes in biological and environmental conditions.

To test our hypothesis, I proposed the following objectives for my research:

1. To study the differences in abundances of proteolytic microbial communities and proteolytic genes under different biotic and environmental conditions.
2. To characterize the microbes involved in protease synthesis in rhizosphere soil and in bulk soil. It is well established as already mentioned that microbes in the rhizosphere soil are affected by rhizodeposition.
3. To determine the shifts in soil protease activity and protease encoding genes due to effect of different plant physiology.

To verify my hypothesis, I have studied:

1. Effect of root exudates and plant varieties on protease activity and protease encoding genes were studied using a rhizobox approach
2. Effect of atmospheric CO_2 on microbial community structure and abundance of protease encoding genes was studied in a free-air CO_2 enrichment (FACE) experiment
3. protease expression under dark and light periods in Barley rhizosphere under green house

References

Adamczyk, B., Godlewski, M., Zimny, J., Zimny, A., (2008). Wheat (*Triticum aestivum*) seedlings secrete proteases from the roots and, after protein addition, grow well on medium without inorganic nitrogen. Plant Biology 10:718–724.

Adamczyk, B., Smolander, A., Kitunen, V., Godlewski, M., (2010). Proteins as nitrogen source for plants. Plant Signal Behav. 2010 Jul; 5(7): 817–819.

Antonov, V.K., Ginodman, L.M., Rumsh, L.D., Kapitannikov, Y.V., Barshevskaya, T.N., Yavashev, L.P., Gurova, A.G., Volkova, L.I., (1981). Studies on the Mechanisms of Action of Protedytic Enzymes Using Heavy Oxygen Exchange. European Journal of Biochemistry, 117, 195-200.

Bach, H. J., Munch, J.C., (2000). Identification of bacterial sources of soil peptidases. Biology and Fertility of Soils, Volume 31,Issue 3, pp 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.

Barrett, A.J., Rawlings, N.D., O'Brien, E.A., (2001). The *MEROPS* Database as a Protease Information System. Journal of Structural Biology 134(2-3),95-102.

Bonmati, M., Ceccanti, B., Nannipieri , P., (1998). Protease extraction from soil by sodium pyrophosphate and chemical characterization of the extracts. Soil Biology and Biochemistry 30(14), 2113–2125.

Brankatschk, R., Töwe, S., Kleineidam, K., Schloter, M., Zeyer, J., (2011). Abundances and potential activities of nitrogen cycling microbial communities along a chronosequence of a glacier forefield. *The ISME* 5, 1025–1037.

Burns, R.G., (1982). Enzyme activity in soil: Location and a possible role in microbial ecology. Soil Biology and Biochemistry, Volume 14, Issue 5, Pages 423-427.

Choudhary, V., Jain, P. C., (2012). Isolation and identification of alkaline protease producing fungi from soils of different habitats of Sagar and Jabalpur District (M.P). Journal of Academia and Industrial Research, 1(3), 106-112.

Das, S.K., Kumar, S., Varma, A., (2010). Role of Enzymes in Maintaining Soil Health. Soil Enzymology, Soil Biology series,Volume 22 , pp 25-42.

Drenth, J., (1980). Proteolytic enzymes: General features of their mode of action, Recueil des Travaux Chimiques des Pays-Bas, Volume 99, Issue 6, pages 185–190, 1980.

Firdaus, R.M., Ahmad, H.A., Sharum, M.Y., Azizi, N., Mohamed, R. (2005). ProLysED: an integrated database and meta-server of bacterial protease systems. Applied Bioinformatics. 4(2):147-50.

Fuka, M.M., Engel, M., Hagn, A., Munch, J.C., Sommer, M., Schloter, M., (2009) Changes of Diversity Pattern of Proteolytic Bacteria over Time and Space in an Agricultural Soil. Microbial Ecology 57:391-401.

Godlewski, M., Adamczyk, B., (2007). The ability of plants to secrete proteases by roots. Plant Physiology and Biochemistry 45:657–664.

Han, W., He, M., (2010). Short-term effects of exogenous protease application on soil fertility with rice straw incorporation. European Journal of Soil Biology, Volume 46, Issue 2, Pages 144–150.

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

Hartley, B., (1960). Proteolytic Enzymes. Annual Review of Biochemistry. 29(1), 45-72.

Hendel, B., Marxsen, J., (2005). Fluorometric determination of the activity of beta-glucosidase and other extracellular hydrolytic enzymes. *Methods to study litter decomposition*. Springer: Dordrecht, pp 261–266.

Hofmockel, K. S., Fierer, N., Colman, B.P., Jackson, R.B., (2010). Amino acid abundance and proteolytic potential in North American soils. Oecologia 163:1069–1078.

Jena, L., Kumar, S., Harinath, B.C., (2012). MycoProtease-DB: Useful resource for *Mycobacterium tuberculosis* complex and nontuberculous mycobacterial proteases. Bioinformation 8(24): 1240-1242.

Kamath, P., Subrahmanyam, V.M., Rao, J.V., Raj, P.V. (2010). Optimization of Cultural Conditions for Protease Production by a Fungal Species. Indian Journal of Pharmaceutical Science ; 72(2): 161–166.

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, Volume 4, Issue 1, Pages 19-30.

Ladd, J.N., Jackson, R.B., (1982) In: Stevenson FJ (ed) Nitrogen in agricultural soils. American Society of Agronomy, WI, pp 173–228.

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

Nannipieri, P., Eldor, P., (2009). The chemical and functional characterization of soil N and its biotic components. Soil Biology and Biochemistry 41 (2009) 2357–2369.

Oseni, O.A. (2011). Production of Microbial Protease from Selected Soil Fungal Isolates. Nigerian Journal of Biotechnology, Vol. 23, 28 – 34.

Pansombat, K., Kanazawa, S., Horiguchi, T. (1997). Microbial Ecology in Tea Soils I. Soil Properties and Microbial Populations. Soil Science and Plant Nutrition 43 (2), 431-438.

Paungfoo-Lonhienne, C., Lonhienne, T.G.A., Rentsch, D., Robinson, N., Christie, M., Webb, R.I., Gamage, H.K., Carroll, B.J., Schenk, P.M., Schmidt, S., (2008).Plants can use protein as nitrogen source without assistance from other organisms. Proceedings of National Academy of Sciences USA 105:4524–4529.

Polgar L. (1989). Mechanisms of Protease Action. CRC Press-Medical - 232 pages

Polgár, L., (2005). The catalytic triad of serine peptidases. Cellular and molecular life sciences, 62, (19-20):2161-72.

Rasche, F., Musyoki, M.K. , Röhl, C., Muema, E.K., Vanlauwe, B., Cadisch, G., (2014).Lasting influence of biochemically contrasting organic inputs on abundance and community structure of total and proteolytic bacteria in tropical soils. Soil Biology and Biochemistry 74, 204-213.

Rawlings, N.D., Barrett, A.J. (1999) MEROPS: the peptidase database. Nucleic Acids Research, 27, 325–331.

Rawlings, N.D., Tolle, D.P., Barrett, A.J., (2004a) MEROPS: the peptidase database. Nucleic Acids Res., 32, D160–D164.

Rawlings,N.D., Tolle,D.P. and Barrett,A.J. (2004b) Evolutionary families of peptidase inhibitors. Biochemical Journal, 378, 705–716.

Rawlings, N.D., Morton, F.R., Barrett, A.J., (2006). MEROPS: the peptidase database. Nucleic Acids Research, 2006, Vol. 34, Database issue, D270–D272.

Rawlings,N.D., Morton,F.R., Kok,C.Y., Kong,J. and Barrett,A.J. (2008) MEROPS: the peptidase database. Nucleic Acids Res., 36, D320–D325.

Rawlings,N.D. (2009) A large and accurate collection of peptidase cleavages in the MEROPS database. Database, doi: 10.1093/ database/bap015.

Rawlings, N.D., Barrett, A.J., Bateman, A., (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Research, Vol. 40, Database issue D343–D350.

Rawlings, N.D., Waller, M., Barrett, A.J., Bateman, A., (2014). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Research, Vol. 42, Database issue D503–D509.

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.

Sakurai, M., Suzuki, K., Onodera, M., Shinano, T., Osaki, M., (2007). Analysis of bacterial communities in soil by PCR–DGGE targeting protease genes. Soil Biology and Biochemistry 39(11), 2777–2784.

Sardans, J., Peñuelas, J., Estiarte, M., (2008). Changes in soil enzymes related to C and N cycle and in soil C and N content under prolonged warming and drought in a Mediterranean shrubland. Applied Soil Ecology, Volume 39, Issue 2, Pages 223–235.

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(2):181-186.

Sharma, A.K., Sharma, V., Saxena, J., Bindu Y., Alam, A., Prakash, A., (2015). Isolation and Screening of Extracellular Protease Enzyme from Bacterial and Fungal Isolates of Soil. International Journal of Scientific Research in Environmental Sciences, 3(9), pp. 0334-0340.

Shun, T., Shigeki, Y., Akio, I., Takayuki, S., Kazuhiro, I. (2014). Linking Temporal Changes in Bacterial Community Structures with the Detection and Phylogenetic Analysis of Neutral Metalloprotease Genes in the Sediments of a Hypereutrophic Lake. Microbes and Environments, Vol. 29, No. 3, 314-321.

Watanabe, K., Hayano, K., (1993) Source of soil protease in paddy fields. Canadian Journal of Microbiology, 39(11): 1035-1040..

Watanabe, K., Hayano, K. (1994). Estimate of the source of soil protease in upland fields, Biology and Fertility of Soils. Volume 18, Issue 4, pp 341-346.

Watanabe, K., Hayano, K., (1994b). 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 (2), Pages 197–203.

Williams, A., (1969). Mechanism of action and specificity of proteolytic enzymes. Quarterly reviews, Chemical Society, 1969, 23, 1-17.

2. Protease encoding microbial communities and protease activity of the rhizosphere and bulk soils of two maize lines with different N uptake efficiency

2.1 Abstract

Present study was carried out to understand the interplay of plant Nitrogen utilizing efficiency (NUE) with potential proteolytic activity and proteolytic community composition of the rhizosphere and bulk soils, sampled from rhizoboxes with two inbred maize lines, L05 and T250, with higher and lower NUE respectively. Microbial biomass was estimated as ATP content and two key bacterial protease encoding genes: alkaline metallo-peptidases (*apr*) and neutral-metallopeptidases (*npr*) were characterized by DGGE and Illumina sequencing of amplicons. Higher protease activity and microbial biomass were observed in rhizosphere soil of the plant line with higher NUE (L05), which also had higher values for Shannon-Weiner diversity indices (H) for DGGE band pattern, with *npr* gene showing higher overall diversity in rhizosphere soil than in the lower NUE plant (T250) rhizosphere. Stronger root effects were observed for *apr* gene than *npr*. Illumina sequencing showed differences in the composition of proteolytic microbial communities in rhizosphere and bulk soils for both L05 and T250, and many unknown *apr* and *npr* gene sequences were also reported. Furthermore, Illumina sequencing results agreed with DGGE data in highlighting higher overall diversity for npr (1,520,600 unique sequences) than for *apr* (934,598 unique sequences). Different members of *Bacillus sp.* were identified as most abundant contributors to *npr* gene pool whereas *apr* gene pool was dominated by genes from *Pseudomonas sp*. This research suggests that plants with different NUE select different bacterial populations with protease encoding genes, which may affect the protease activity of the rhizosphere soil.

2.2 Introduction

Genetic and physiological mechanisms of N acquisition by important cereal plants are increasingly known (Hirel et al., 2007) but currently, at field scale, the Nitrogen Use Efficiency (NUE) in cereal production is still lower than 40% (Raun and Johnson, 1999). This is because although the NUE is an inherent plant characteristic, regulated by complex genetic and metabolic factors (Xu et al. 2012, Ngezimana and Agenbag, 2014; Zamboni et al., 2015), the N acquisition by crop plants is also limited by N losses by volatilization, runoff and leaching, and by microbial N immobilization. Moreover, there are increasing evidences that plant NUE also depends on microbial activity in the rhizosphere, particularly on activity of the proteolytic communities (Mooshammer et al. 2004). This is linked to the fact that the most of soil N is of peptidic or protein origin, as 96-99% of soil total N is organic and after acid hydrolysis, amino acidic N accounts for 30-50% of the N in soil (Nannipieri and Paul, 2009). The N phytoavailability in soil also depends on the hydrolysis of other organic N forms, such as urea and chitin catalyzed by the urease and chitinases, the latter being produced by fungi and bacteria (Metcalfe et al., 2002), Chitinase is, therefore, a key soil enzyme, regulating the release of low molecular weight N-sugars from which N is rapidly mineralized to inorganic N (Gooday 1994). Proteins in soil originate from plants, animals and microorganisms, either through active excretion or passive release, and therefore a high proportion of protein N in the rhizosphere is expected. In soil environment, protein N is released after protein hydrolysis by extracellular proteases of plants, animals and microbial origin (Adamczyk et al., 2010, Godlewski and Adamczyk, 2007; Hayano 1993; Watanabe, 2009), and previous studies indicated that metalloproteases of bacterial origin mainly contribute to the measured soil protease activity (Hayano et al. 1987, Bach and Munch, 2000, Kammimura and Hayano, 2000).

Soil management and environmental factors influence the abundance and distribution of microbial genes encoding for neutral metallo-peptidases (*npr*), alkaline metallo-peptidases (*apr*) and serine peptidases (*sub*) (Bach et al. 2001, 2002, Fuka et al. 2008a, Rasche et al. 2014; Fuka et al. 2008, 2009; Sakurai et al. 2007; Tsuboi et al. 2014, Fuka et al 2009). Proteases catalyze the hydrolysis of the terminal amino acid of polypeptide chains (exopeptidases) or of internal peptide bond (endopeptidases) on one or few related substrates, with the majority of proteases acting on several substrates. However, the number of assays for soil protease activity is limited to few substrates and optimal pH values. Increased N mineralizing activities in response to the release of root exudates has been reported (Renella et al., 2007), but in spite of their importance in determining N availability to plants, studies on the link between the diversity of protease encoding genes and protease activities in the rhizosphere are still scarce (Nannipieri et al., 2012). Little information is also available on the relations between the proteolytic microbial community of the rhizosphere and the plant NUE. Next generation sequencing technologies (NGS) provide advanced tools to analyze microbial genes in soil: this approach has been applied for the analyses of PCR amplicons of 16S rRNA (Vasileiadis et al. 2013), ITS (Internal Transcribed Spacer) (McHugh and Schwartz, 2015) and ammonia monooxygenases (Pester et al., 2012), but not yet for assessing the abundance and diversity of proteases genes in soil.

We hypothesized that plants with different NUE select different proteolytic microbial communities characterized by different levels of proteolytic activity in the rhizosphere. To test our hypotheses, we studied the composition of the proteolytic microbial communities and proteolytic activities in the rhizosphere and bulk soil of the L05 and T250 maize lines, characterized by high and low NUE, respectively. Previous work showed that these two

maize lines have different genetic responses to N availability (Zamboni et al., 2014) and also host different microbial communities in their rhizosphere (Pathan et al., 2015). Furthermore, we have also applied a NGS assessment of neutral metallo-peptidases (*npr*) and alkaline metallo-peptidases (*apr*) PCR amplicons, in order to unravel the diversity of these genes in the bulk and rhizosphere soils of the two maize lines. We also measured the urease and chitinase activities to understand their contribution to N availability in the maize rhizosphere. Results of this research can improve our understanding of the effects of microbial selection in the rhizosphere of maize plants with different plant NUE on the turnover of protein-N in the rhizosphere.

2.3 Materials and methods

Soil properties and rhizobox set up

A sandy clay loam Eutric Cambisol (World Reference Base for Soil Resources, 2006), under conventional maize crop regime, located at Cesa (Tuscany, Central Italy), was sampled from the Ap horizon (0-25 cm). Soil had a pH value (in H_2O) of 7.1, contained 32.1% sand, 42.2% silt, 25.7% clay, 10.8 g kg^{-1} total organic C (TOC), 1.12 g kg^{-1} total N and 6.45 g kg^{-1} total P. The soil was sieved at field moisture (< 2 mm), after removing visible plant material. After sieving, 600 g of soil were placed in the soil compartment of the rhizoboxes as reported by Pathan et al. (2015). The L05 and T250 maize lines were grown for 21 and 28 d, respectively, a suitable growth period to allow the full colonization of the plant compartment by plant roots and prevent nutrient starvation. Plants were regularly watered with distilled sterile H_2O and no fertilizers were applied during the plant growth. Full details on the maize growth conditions were reported by Pathan et al. (2015). Five rhizoboxes replicates for each maize line were prepared. The used rhizoboxes allowed precise sampling of rhizosphere due to the presence of fixed sampling groves at precise increment distances from the surface of the plant compartment. Rhizosphere (R) and bulk soil (B) samples of the L05 and T250 maize lines were

named as L05 R, L05 B, T250 R and T250 B, respectively. Rhizosphere and bulk samples were kept separate after sampling, and immediately analyzed for the enzyme activities or stored at -80°C before ATP determination or DNA extraction.

Soil microbial biomass and enzymatic activities

Soil microbial biomass was estimated by determining the ATP content according to Ciardi and Nannipieri (1990). The N-benzoyl-L-argininamide (BAA) and casein hydrolyzing activities (protease activities) were determined according to Ladd and Butler (1972) and Nannipieri et al (1974), respectively. Concentrations of NH_4^+-N and tyrosine released by the assays with BAA or Na-caseinate, respectively, were spectrophotometrically quantified (Perkin Elmer Lambda 2) from calibration curves obtained using standards after reaction with the Nessler or Folin reagents after subtracting of the absorbance of controls. Urease activity was determined using 6% urea solution as substrate according to Nannipieri et al. (1980), and NH_4^+-N concentration was determined as above described for the for the protease assay. To account for fixation of NH_4^+-N released by BAAase and urease activities, NH_4^+-N solutions with concentrations in the range of those released by urease and protease activities were incubated with the same soil, and recovery of NH_4^+-N were in the range 95-98%. Chitinase activity was determined by the hydrolysis of 4-nitrophenyl-β-D-glucosaminidine (SIGMA) in 0.1 acetate buffer at pH 5.2, for 1 h at 50°C using 1 g d.w. soil. The p-nitrophenol released by the chitinase activity was spectrophotometrically quantified using calibration curves, after subtracting of the absorbance of controls. Preliminary experiments have showed that 50°C was the optimal temperature for soil chitinase activity.

Nucleic acids extraction and PCR-DGGE analysis

DNA was extracted by sequential extraction method from 0.5 g soil as described by Ascher et al. (2009) using the FastDNA spin kit for soil (MP Biomedicals, USA), and the intracellular DNA fraction was used in this study. The DNA yield and purity were analysed with a Qubit 2.0 fluorometer (Life

Technologies, USA) using Quant-iT dsDNA *HS* kit according to the manufacturer's instructions, and stored at -20°C till prior to analysis.

The primers F*apr*I/R*apr*II for *apr* gene and FnprI/RnprII for *npr* as mentioned in Bach et al. (2001) were used for PCR and were amplified according to conditions as used by Bach et al. (2001). The DGGE conditions for the fingerprinting of the *apr* and *npr* amplicons were those previously used bySakurai et al. (2007). The DGGE fingerprints were performed using a INGENY PhorU System (Ingeny International BV, Netherlands), the DGGE gels were stained with SybrGreen I (FMC Bio Products, Rockland, ME, USA), and the banding patterns were analysed by a Gel Doc system (Bio-Rad, USA).

Quantification of protease encoding genes

Quantitative PCR (qPCR) was conducted on a CFX Connect Real-Time PCR Detection System (Bio-rad Laboratories) to determine the abundance of *apr* and *npr* genes, using the primer sets FP *apr*I/ RP *apr*II for the *apr* gene and FP *npr*I/ RP *npr*II for the *npr* gene, according to Bach et al. (2001). *Pseudomonas fluorescence* (isolated from an agricultural soil) and *Bacillus cereus* (DSM31) were used as positive controls for *apr* and *npr* genes respectively. Each qPCR assay was conducted in a 96-well plate and included three replicates for each standard, negative controls, and sample. Amplification was performed using the iTaq Universal SYBR Green Supermix (Bio-rad Laboratories), adding to each reaction mixture forward and reverse primers for both genes at concentration of 0.6 µM, 3% of bovine serum albumin(BSA), 20 ng DNA template for *apr* gene and 30 ng DNA template for *npr*. The PCR runs for both genes started with an enzyme activation step at 95°C for 3 min, followed by 42 cycles of denaturation at 94°C for 25 s. Annealing conditions were 54°C for 30 s for the *apr* and at 53°C for 30 s for the *npr* gene, respectively, followed by extension at 72°C for 30 s. The specificity of amplification products were confirmed by melting curve analysis and expected sizes of amplified fragments were checked by running the amplicons on a 2% agarose gel stained with ethidium bromide for 90 mins at 100 V.

Illumina sequencing of *apr* and *npr* genes

Alkaline metallo-peptidase (*apr*) and neutral metallopeptidase (*npr*) genes were targeted by PCR as previously described (Bach et al., 2001), using primers pairs FP *apr*I/RP *apr*II for *apr* (amplicon length 194 bp) and FP nprI/RP nprII for the *npr* gene (amplicon length 233 bp), respectively. The PCR reactions were carried out on a Biometra T Professional thermocycler (*Biometra* Biomedizinische Analytik GmbH, Germany). For both *apr* and *npr* genes the reaction mixture contained 0.8 µM of forward and reverse primers, 20 ng of template DNA, 0.3% BSA, 0.2 mM dNTP mix, 2.5 µl of 10X DreamTaq Buffer having 20 mM $MgCl_2$ and 1 unit of Dream Taq Polymerase (Thermo Fisher Scientific, USA). The PCR programs consisted of a hot start step for 5 min at 95°C, followed by 80 °C for 5 min during which Taq polymerase was added. Thirty-five cycles of denaturation at 94°C for 30 s, annealing at 55°C for *npr* and 58°C for *apr*, respectively, followed by an extension step at 72°C for 30s and a final extension step at 72°C for 7 mins. After PCR, amplicons were run on a 2% agaorose gel for 90 mins, single bands were excised and purified from gel using Nucleospin Gel and PCR cleanup kit (MACHERY-NAGEL GmbH and Co. KG, Germany), according to the manufacturer's instructions. Purified amplicons were quantified on Qubit 2.0 fluorometer using Quant-iT dsDNA *HS* reagent as per manufacturer's instructions and sequenced using an Illumina HiSeq 2000 in paired-end 150 x 2 bp at the Beijing Genomics Institute. For Illumina sequencing the five replicates of each plant were pooled together for an in depth analysis of all gene sequences, according to the Illumina sample preparation guide (http://supportres.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf). The obtained Illumina sequences of *apr* and *npr* amplicons were processed separately, using the USEARCH and UPARSE pipelines (Edgar, 2010, 2013). Paired MiSeq reads from each sample were firstly assembled with the fastq_mergepairs command. Assembled reads were then filtered allowing a maximum expected error of 0.5 and

discarding reads with length <190 and < 230 bp for *apr* and *npr* gene, respectively. Barcode labels were then added to the sequences, and 4 samples available for each gene were merged with the cat command. Unique sequences were then identified, sorted by abundance, and singletons were discarded. Gene sequences were assigned to operational taxonomical units (OTUs) at minimum identity levels of 97% or 95%, according to the UPARSE algorithm. The OTUs were further filtered for the presence of chimeras with the UCHIME tool, and reads were finally mapped back to obtain OTUs abundance. For each gene, sequences were pooled together and dereplicated in order to identify and count the unique sequences.

Data analyses

Microbial biomass and enzyme activities data were analyzed by ANOVA. The significance of differences between mean values were determined by the Fisher PLSD. For PCR-DGGE analysis, bands were identified and their intensities were measured after normalizing lanes and background subtraction using Quantity-One® software (Bio-Rad Laboratories, USA). Band intensities were used to calculate the Shannon-Weaver diversity index H (Shannon and Weaver, 1963) according to the eq. 1, using the PAST software (Hammer et al. 2001),

$$H = -\sum \left(\frac{n_i}{N}\right) \log(ni/N) \qquad \text{(eq. 1)}$$

where n_i is the relative intensity of each DGGE band, S is the number of DGGE bands for each lane and N is the sum of intensities for all bands in a given sample (or lane). The DGGE banding pattern was clustered to UPGAMA dendrograms based on Raup and Crick similarity indices (Raup and Crick, 1979) using the PAST software. A principal component analysis (PCA) for enzyme activity data and Shannon-Wiener diversity index were carried out based on correlation matrix and results were displayed as biplot using PAST.

Analysis of the Illumina sequencing data

Mothur v. 1.32.1 was used for calculating diversity indexes and rarefaction curves from the OTU data (Schloss et al., 2009). The OTUs fasta sequences were analysed and annotated on NCBI with blastx and blastn using the Blast2go software (Conesa et al., 2005). Phylogenetic trees were constructed on the aligned sequences with the PhyML (Phylogeny Maximum Likelihood) approach (Guindon and Gascuel, 2003) by applying the Shimodaira–Hasegawa [SH]-aLRT test, and alignments and tree generation were carried out using the SeaView software (Gouy et al., 2010).

2.4 Results

Microbial biomass and enzyme activities

Microbial biomass based on ATP content was significantly higher in the rhizosphere of the L05 maize line, as compared to its bulk soil, whereas no significant differences were observed between rhizosphere and bulk soil of the T250 maize line (Figure 1a). BAA hydrolyzing activity was significantly higher in the rhizosphere of both L05 and T250 maize line, as compared to their respective bulk soils (Figure 1b). Caseinase hydrolyzing activity was only enzyme activity that was significantly lower in the rhizosphere of both L05 and T250 maize line, as compared to their respective bulk soils (Figure 1c), and also it was only enzyme activity that was significantly lower in the rhizosphere of the L05 than in the T250 maize line rhizosphere. Chitinase activity was significantly higher in the rhizosphere of the L05 than its respective bulk soil, whereas there was no significant difference between rhizosphere and bulk soil of the T250 maize line (Figure 1d). Moreover, the chitinase activity was significantly higher in the rhizosphere of the L05 than in the T250 maize line rhizosphere (Figure 1d). Urease activity was significantly higher in the rhizosphere of both L05 and T250 maize line, as compared to their respective bulk soils (Figure 1E). Moreover, the urease activity was significantly higher in the rhizosphere of the L05 than in the T250 maize line rhizosphere (Figure 1d). The PCA analysis showed that ATP, Urease, BAAase hydrolysing and

chitinase activities were related to each other, but not related to Caesinase hydrolysing activity (Figure 7).

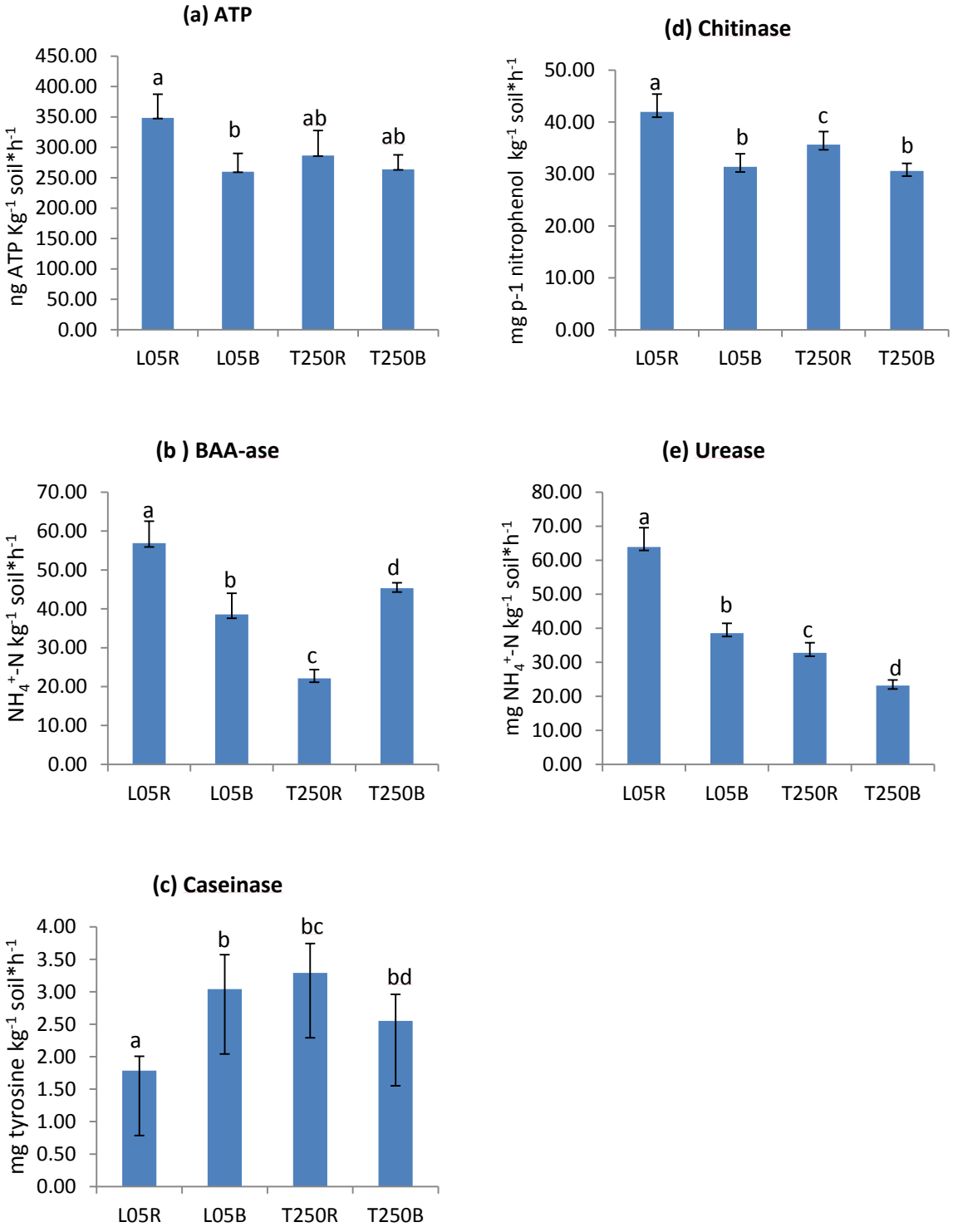


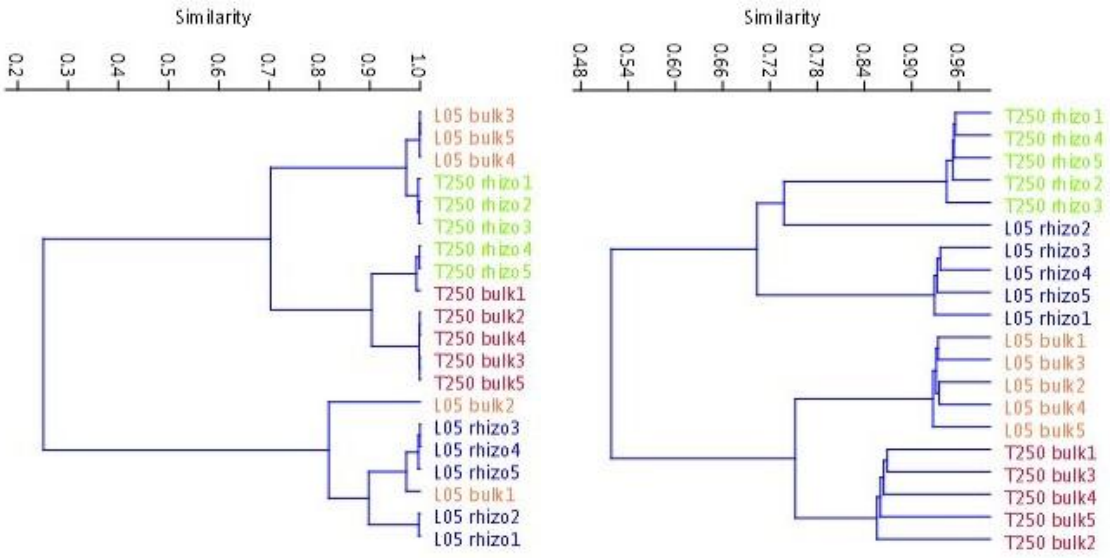
Figure 1. Activity results (a) ATP content, (b) BAA-ase activity, (c) caseinase activity, (d) chitinase activity, (e) urease activity of the rhizosphere and bulk soil of the Lo5 and T250 maize lines. Values are the mean of five replicates and the error bars represent the standard deviation of the mean values.

PCR-DGGE microbial community composition

The DGGE analysis showed complex banding patterns for both *npr* and *apr* genes. The UPGAMA based on Raup and Crick's similarity index for *npr* and *apr* genes showed that the rhizosphere and bulk soils of the L05 and T250 maize lines clustered separately, although the separation between clusters was not significant (Figure 2).The Shannon-Wiener diversity indices for the *apr* gene showed a significantly greater ($P < 0.05$) diversity in the rhizosphere of both maize lines, as compared to their respective bulk soils: the diversity indices for the *apr* gene could be ranked as : T250B, > L05R > L05B > T250R. (Table 1) and *npr* gene diversity didn't show rhizosphere effect.

	Shannon-Weiner diversity index (H) and standard deviation			
	apr gene		npr gene	
L05R	1.359 [a]	±0.006	2.032 [a]	±0.018
L05B	1.328 [a]	±0.020	1.729 [b]	±0.009
T250R	0.691 [b]	±0.001	2.040[a]	±0.004
T250B	1.554 [c]	±0.021	1.919 [c]	±0.003

Table 1. Values of the Shannon-Weiner indices for DGGE bands for the npr and *apr* genes in the rhizosphere and bulk soil of the L05 and T250 maize lines. Values are shown as mean (n = 5) and standard deviation, and different superscripts indicate significant differences ($P < 0.05$) of values within each column.



2(a) 2 (b)

Figure 2. UPGAMA custers based on Raup Crick similarity for (a) npr gene and (b) *apr* gene

Protease gene quantification

The qPCR analysis showed a significantly ($P < 0.05$) higher number of *apr gene* copies in the rhizosphere and bulk soil of the L05 as compared to the T250 maize line, whereas for the *npr* gene there were no significant differences between the copy numbers regardless of the maize line and soil type (Table 2). The PCA on qPCR, gene diversity, ATP data and enzyme activities showed that the rhizosphere of the high NUE L05 maize line clustered separately from the respective bulk soil and from the T250 rhizosphere and bulk soil (Figure 7). The PCA also showed that both *apr* and *npr* gene abundances clustered together, with higher correspondence to the BAA-hydrolyzing activity than to the caseinase hydrolyzing activity (Figure 7).

Soil type	Copy numbers and standard deviation			
	apr gene		npr gene	
L05R	$2.7 \cdot 10^{5\,a}$	$\pm 1.5 \cdot 10^{5}$	$7.1 \cdot 10^{7\,a}$	$\pm 1.1 \cdot 10^{7}$
L05B	$1.8 \cdot 10^{5\,a}$	$\pm 1.4 \cdot 10^{5}$	$5.9 \cdot 10^{7\,a}$	$\pm 1.1 \cdot 10^{7}$
T250R	$1.2 \cdot 10^{5\,b}$	$\pm 8.8 \cdot 10^{4}$	$4.2 \cdot 10^{7\,a}$	$\pm 1.2 \cdot 10^{7}$
T250B	$1.4 \cdot 10^{5\,b}$	$\pm 6.5 \cdot 10^{4}$	$3.0 \cdot 10^{7\,b}$	$\pm 5.3 \cdot 10^{6}$

Table 2 Gene copy numbers per gram of soil for npr and *apr* genes in the rhizosphere and bulk soil of the L05 and T250 maize lines. Values are shown as mean (n = 5) and standard deviation, and different superscripts indicate significant differences ($P < 0.05$) of values within each column.

Protease high throughput sequencing analyses

Assembly of paired-reads was correctly performed for more than 99% of sequences for each sample (Table 3). Rarefaction curves indicated a very good coverage for both *apr* (Figure 3a) and *npr* (Figure 3b) genes, as confirmed by Good's coverage values that were always higher than 99.99% (Table 4). Rarefaction curves indeed show that approximately 100,000 sequences may be enough to get a good picture of *apr* and *npr* genes diversity in agricultural soils. After discarding ambiguous sequences and sequences shorter than target length for *apr* gene (190 bp) and *npr* (230 bp) amplicons, the retained sequences were 49.2 % and 72.4 % for the *apr* gene and *npr* genes, respectively.

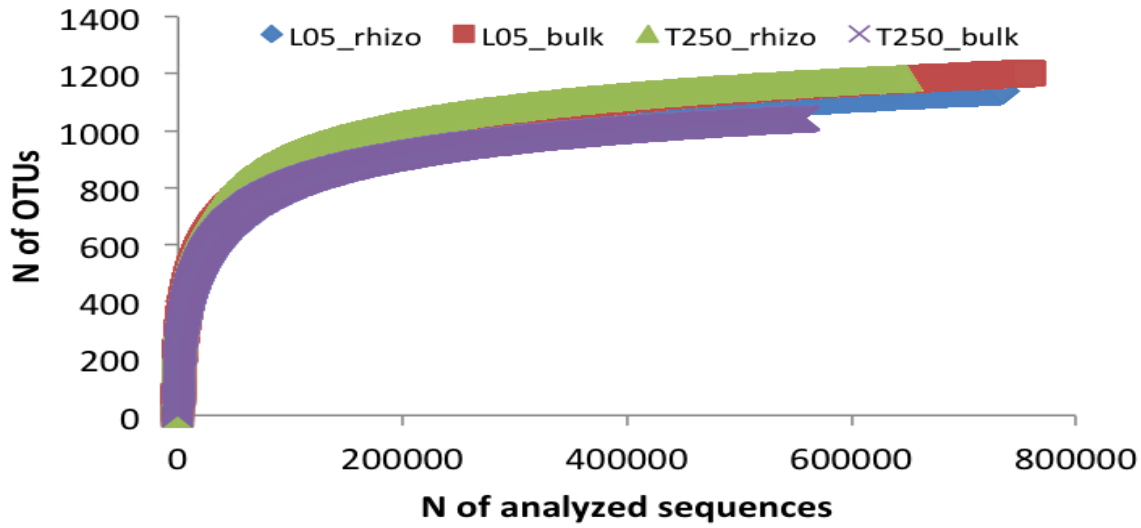


Figure 3 (a)

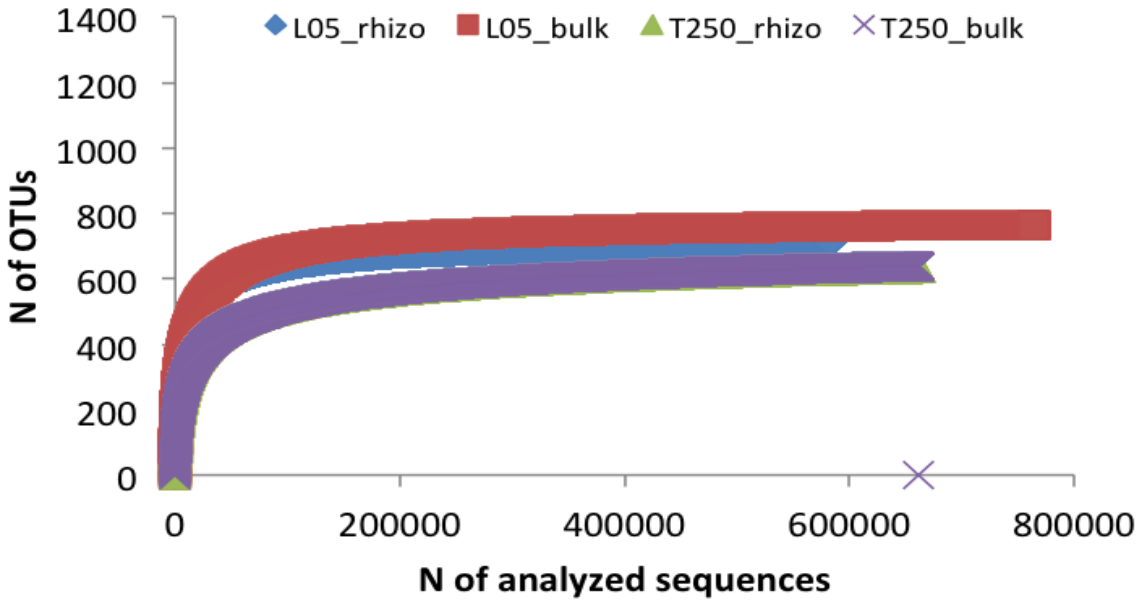


Figure 3 (b)

Figure 3. Rarefaction curves for of *apr* (a) and *npr* (b) amplicons sequences generated in Illumina.

Label	Gene	Sample	n of paired sequences assembled (%)	n of sequences passing filter[a]	total n of unique sequences	final n of OTUs at 97%	final n of OTUs at 95%
D1	*apr*	L05_rhizo	1,435,410 (99.8%)	867,314 (60.4%)	934,598	1763	1136
D2	*apr*	L05_bulk	1,215,724 (100%)	879,698 (72.4%)		1844	1201
D3	*apr*	T250_rhizo	1,110,522 (99.9%)	766,169 (69.3%)		1797	1183
D4	*apr*	T250_bulk	1,185,430 (99.9%)	677,685 (57.2%)		1664	1041
D5	npr	L05_rhizo	1,480,616 (99.8%)	728,431 (49.2%)	1,520,600	1331	712
D6	npr	L05_bulk	1,367,344 (99.8%)	938,675 (68.6%)		1421	765
D7	npr	T250_rhizo	1,399,313 (99.9%)	806,304(57.6%)		1242	631
D8	npr	T250_bulk	1,185,430 (99.9%)	804,063 (63.7%)		1239	639

[a] maximum error 0.5, length > 190 bp for *apr*, > 230 bp for npr

Table 3. Preprocessing and OTUs clustering of *apr* and *npr* Illumina reads.

Label	Gene	Sample	Coverage	Simpson eveness	Inverted Simpson	Chao	Shannon eveness	non parametric Shannon
D1	*apr*	L05_rhizo	99.99%	0.039	44.66	1292.3	0.645	4.53
D2	*apr*	L05_bulk	99.99%	0.031	37.82	1318.7	0.626	4.44
D3	*apr*	T250_rhizo	99.99%	0.036	43.45	1297.1	0.649	4.59
D4	*apr*	T250_bulk	99.99%	0.034	35.99	1117.0	0.634	4.41
D5	npr	L05_rhizo	99.99%	0.028	19.95	727.0	0.609	4.01
D6	npr	L05_bulk	99.99%	0.025	18.96	776.3	0.597	3.97
D7	npr	T250_rhizo	99.99%	0.022	14.02	658.4	0.536	3.46
D8	npr	T250_bulk	99.99%	0.022	14.71	679.6	0.542	3.51

Table 4. Coverage, diversity and richness indexes in the analyzed *apr* and *npr* Illumina reads.

A total of 9,34,598 and 1,520,600 unique sequences were obtained for the *apr* and for the *npr* gene respectively (Table 3). Clustering of these sequences at 97% similarity resulted in

1767 and 1308 average OTUs for the *apr* and *npr* gene, respectively (Table 3). Blastx results at 97% identity showed that many OTUs gave the same hits, albeit their nucleotidic sequences were different; for this reason analyses were also performed with OTUs at 95% similarity, in this case the number of detected OTUs per sample were as expected lower, varying between 631 for T250 rhizosphere and 765 for L05 bulk sample (Table 3). Results herewith presented refer to the analyses of OTUs with 95% minimum identity. Analysis of the OTUs was conducted on the first 50 most abundant OTUs covering 74% of total OTUs diversity for *apr* and 85.4% for *npr* (Tables 5 and 6).

Table 5. Functional annotation of the first 50 most abundant *apr* OTUs, covering average of 74.7% of total diversity. Percentage per sample for each OTU is reported, together with the results of MEROPS annotation.

OTUId	L05 rhizo	L05 bulk	T250 rhizo	T250 bulk	MEROPS code	Annotation with best hit
OTU_1	7.2	5.7	7.1	8.1	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_3	5.8	5.7	7.1	8.1	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_2	2.6	6.7	2.7	1.7	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_8	3.1	2.8	3.6	4.4	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_6	2.9	2.8	2.9	4.6	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_12	2.0	6.4	2.2	1.5	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_4	3.3	2.3	3.4	2.2	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)

OTU_13	2.0	2.3	2.9	4.0	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_32	1.9	2.0	2.8	3.6	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_10	2.3	1.9	3.3	2.1	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_7	2.9	1.7	2.5	2.5	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_24	2.3	1.8	1.9	3.0	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_25	2.3	1.6	2.6	2.5	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_5	1.4	2.2	2.5	2.5	MER301609	streptogrisin B (Streptomyces griseus).
OTU_418	1.8	1.6	1.9	2.6	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_9	4.9	1.5	0.0	0.2	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_23	0.9	1.6	2.1	2.0	MER599609	family I63 unassigned peptidase inhibitors (Echinops telfairi)
OTU_52	1.3	1.3	1.4	2.1	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_15	1.3	1.2	1.8	1.0	MER611137	family I43 unassigned peptidase inhibitors
OTU_11	0.1	4.5	0.0	0.0	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_14	0.1	4.4	0.0	0.0	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_17	1.6	0.8	1.4	0.4	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_27	1.3	0.8	1.5	0.5	MER005073	subfamily M10B unassigned

						peptidases (Proteus mirabilis)
OTU_727	1.3	2.0	0.3	0.1	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_20	0.9	0.8	1.4	0.7	MER362494	FtsH peptidase ({Thermotoga}-type) (Nitrosococcus watsoni)
OTU_19	2.0	0.8	0.0	0.1	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_1484	0.8	0.6	0.8	0.9	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_604	1.3	0.3	0.9	0.3	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_29	2.4	0.1	0.0	0.0	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_1322	1.6	0.7	0.0	0.1	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_388	0.6	0.4	0.6	1.1	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_854	0.7	0.4	0.5	1.1	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_33	0.9	0.3	0.5	0.7	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_331	0.8	0.3	0.7	0.4	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_650	0.2	1.7	0.0	0.0	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_325	0.4	0.4	0.6	0.8	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_151	0.6	0.4	0.5	0.7	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_16	0.3	0.2	0.8	1.0	MER127948	subfamily M24A unassigned peptidases (Beutenbergia cavernae)

OTU_785	0.8	0.3	0.8	0.2	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_44	0.7	0.3	0.5	0.7	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_61	0.2	0.0	0.7	1.3	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_48	0.3	0.3	0.8	0.6	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_38	0.2	0.2	0.9	0.6	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_534	0.6	0.2	0.7	0.3	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_26	0.2	0.2	0.7	0.8	MER273427	subfamily S26B unassigned peptidases (Paenibacillus mucilaginosus)
OTU_108	0.2	0.2	0.5	1.0	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_18	0.5	0.4	0.3	0.5	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_34	0.3	0.3	0.6	0.5	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_53	0.4	0.4	0.5	0.3	MER005073	subfamily M10B unassigned peptidases (Proteus mirabilis)
OTU_22	0.6	0.3	0.3	0.3	MER229708	repressor LexA (Bacillus cellulosilyticus)
Cumul %	75.1	76.3	72.8	74.6		

Table 6. Functional annotation of the first 50 most abundant *npr* OTUs, covering on average the 85.4% of total diversity. Percentage per sample for each OUT is reported, together with the results of MEROPS annotation.

OTUId	L05 rhizo	L05 bulk	T250 rhizo	T250 bulk	MEROPS code	Annotation with best hit
OTU_1	12.2	11.9	11.8	11.9	MER001031	thermolysin (Bacillus megaterium)
OTU_5	8.7	10.2	13.6	13.2	MER109427	stearolysin (Bacillus sp. SG-1)
OTU_3	11.1	10.6	10.9	10.8	MER001031	thermolysin (Bacillus megaterium)
OTU_6	7.8	9.0	12.2	11.3	MER109427	stearolysin (Bacillus sp. SG-1)
OTU_9	5.3	5.5	6.7	6.5	MER038281	stearolysin (Bacillus vietnamensis)
OTU_10	5.0	5.0	6.3	5.9	MER038281	stearolysin (Bacillus vietnamensis)
OTU_11	2.5	2.5	2.8	2.9	MER038281	stearolysin (Bacillus vietnamensis)
OTU_13	2.5	2.2	2.7	2.7	MER038281	stearolysin (Bacillus vietnamensis)
OTU_17	2.1	2.1	2.5	2.8	MER109427	stearolysin (Bacillus sp. SG-1)
OTU_19	2.2	2.1	2.4	2.6	MER109427	stearolysin (Bacillus sp. SG-1)
OTU_27	1.0	1.0	1.3	1.3	MER014941	stearolysin (Clostridium acetobutylicum)
OTU_23	0.9	0.8	1.2	1.1	MER014941	stearolysin (Clostridium acetobutylicum)
OTU_2	0.9	1.0	0.8	1.1	MER003136	immunoproteasome catalytic subunit 3 (Sus scrofa)
OTU_8	0.9	1.0	0.7	1.0	MER169972	amblin inhibitor unit 2 ({Amblyomma hebraeum})
OTU_7	0.9	0.7	0.7	0.9	MER253284	subfamily M15C unassigned peptidases (Spirochaeta caldaria)
OTU_4	0.9	0.7	0.7	0.8	MER025290	subfamily C14B unassigned peptidases (Streptomyces coelicolor)
OTU_28	0.6	0.5	1.0	0.9	MER224756	stearolysin (Brevibacillus agri)

OTU_18	1.4	0.7	0.2	0.6	MER001026	thermolysin (Bacillus thermoproteolyticus)
OTU_45	0.5	0.4	0.9	0.7	MER224756	stearolysin (Brevibacillus agri)
OTU_22	0.8	0.6	0.6	0.5	MER001031	thermolysin (Bacillus megaterium)
OTU_15	0.7	0.6	0.6	0.5	MER001031	thermolysin (Bacillus megaterium)
OTU_25	1.2	0.6	0.2	0.5	MER001026	thermolysin (Bacillus thermoproteolyticus)
OTU_20	0.6	0.5	0.5	0.7	MER187790	stearolysin (Bacillus pseudomycoides)
OTU_48	0.8	0.4	0.6	0.4	MER029719	neutral peptidase ({Thermoactinomyces}-type) (Thermoactinomyces sp. 27a)
OTU_203	0.5	0.4	0.7	0.6	MER038281	stearolysin (Bacillus vietnamensis)
OTU_29	0.5	0.5	0.5	0.6	MER187790	stearolysin (Bacillus pseudomycoides)
OTU_26	0.7	0.4	0.4	0.5	MER224756	stearolysin (Brevibacillus agri)
OTU_185	0.4	0.4	0.6	0.5	MER038281	stearolysin (Bacillus vietnamensis)
OTU_42	0.6	0.4	0.4	0.5	MER224756	stearolysin (Brevibacillus agri)
OTU_652	0.5	0.4	0.5	0.4	MER109427	stearolysin (Bacillus sp. SG-1)
OTU_38	0.6	0.6	0.3	0.2	MER001031	thermolysin (Bacillus megaterium)
OTU_115	0.5	0.6	0.3	0.4	MER001031	thermolysin (Bacillus megaterium)
OTU_61	0.6	0.3	0.5	0.3	MER029719	neutral peptidase ({Thermoactinomyces}-type) (Thermoactinomyces sp. 27a)
OTU_92	0.5	0.5	0.3	0.4	MER001031	thermolysin (Bacillus megaterium)
OTU_41	0.6	0.5	0.2	0.2	MER001031	thermolysin (Bacillus megaterium)
OTU_74	0.4	0.4	0.3	0.2	MER029719	neutral peptidase ({Thermoactinomyces}-type) (Thermoactinomyces sp. 27a)
OTU_49	0.5	0.3	0.2	0.3	MER001354	thermolysin (Lactobacillus sp.)

OTU_56	0.3	0.3	0.2	0.2	MER029719	neutral peptidase ({Thermoactinomyces}-type) (Thermoactinomyces sp. 27a)
OTU_60	0.3	0.3	0.2	0.3	MER109364	stearolysin (Bacillus sp. SG-1)
OTU_73	0.5	0.5	0.0	0.0	MER287706	stearolysin (Planococcus donghaensis)
OTU_55	0.4	0.3	0.2	0.2	MER029719	neutral peptidase ({Thermoactinomyces}-type) (Thermoactinomyces sp. 27a)
OTU_31	0.2	0.3	0.2	0.3	MER109427	stearolysin (Bacillus sp. SG-1)
OTU_46	0.3	0.3	0.2	0.3	MER311923	subfamily M15B non-peptidase homologues (Streptococcus intermedius)
OTU_51	0.0	0.9	0.0	0.0	MER001031	thermolysin (Bacillus megaterium)
OTU_77	0.3	0.3	0.2	0.2	MER109364	stearolysin (Bacillus sp. SG-1)
OTU_39	0.2	0.3	0.3	0.3	MER001030	thermolysin (Bacillus cereus)
OTU_53	0.2	0.3	0.2	0.3	MER001030	thermolysin (Bacillus cereus)
OTU_14	0.1	0.1	0.4	0.3	MER253096	FtsH peptidase ({Thermotoga}-type) (Nitrosomonas sp. Is79A3)
OTU_88	0.4	0.2	0.2	0.2	MER001354	thermolysin (Lactobacillus sp.)
OTU_58	0.5	0.4	0.0	0.0	MER287706	stearolysin (Planococcus donghaensis)
Cumul %	82.2	80.7	89.5	89.2		

The most abundant *apr* OTUs revealed high phylogenetic similarity with *Pseudomonas* sp, followed by *Caulobacter* sp. and *Dickeya* sp. (Figures 4a and 4b).

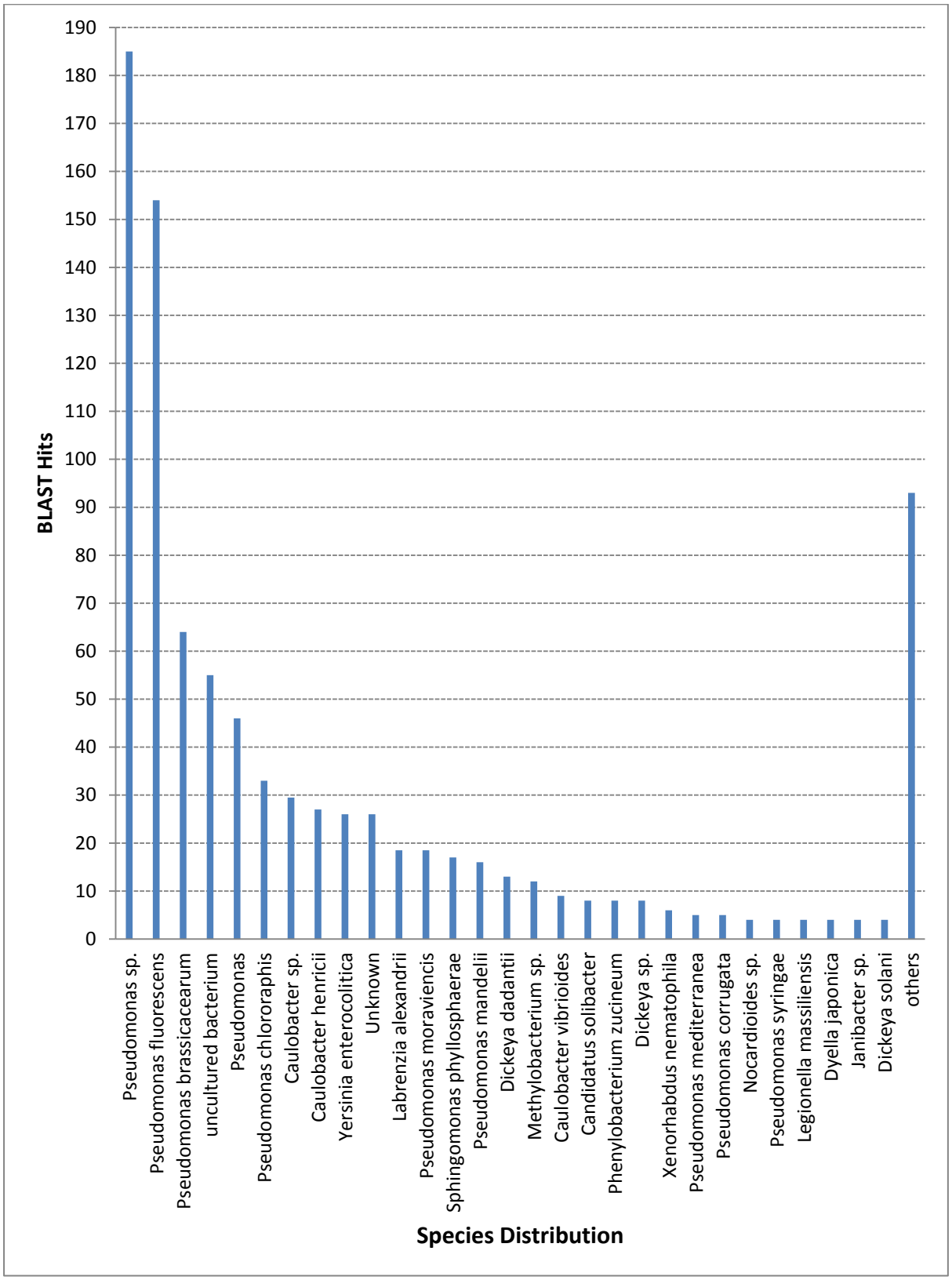


Figure 4 (a)

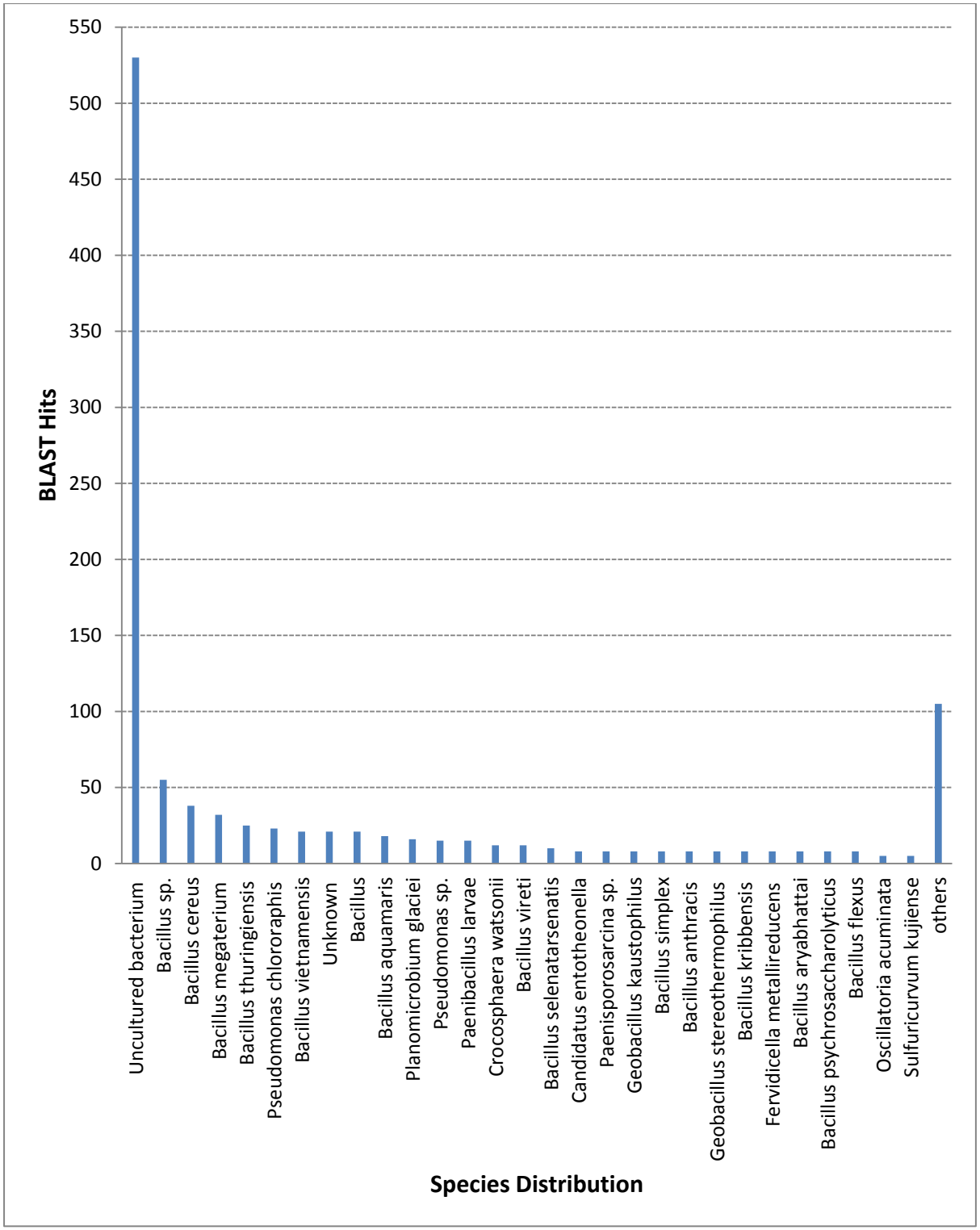


Figure 4 (b)

Figure 4 Species distribution of the hits of the 50 most abundant (a) *apr* OTUs and (b) npr OTUs

Both hierarchical clustering (Figures 5a and 5b) and PCA (Figures 6a and 6b) analyses indicated that the protease gene diversity was influenced by the maize line, less from the rhizosphere or bulk soil, particularly for the T250 maize line (Figures 6a and 6b).

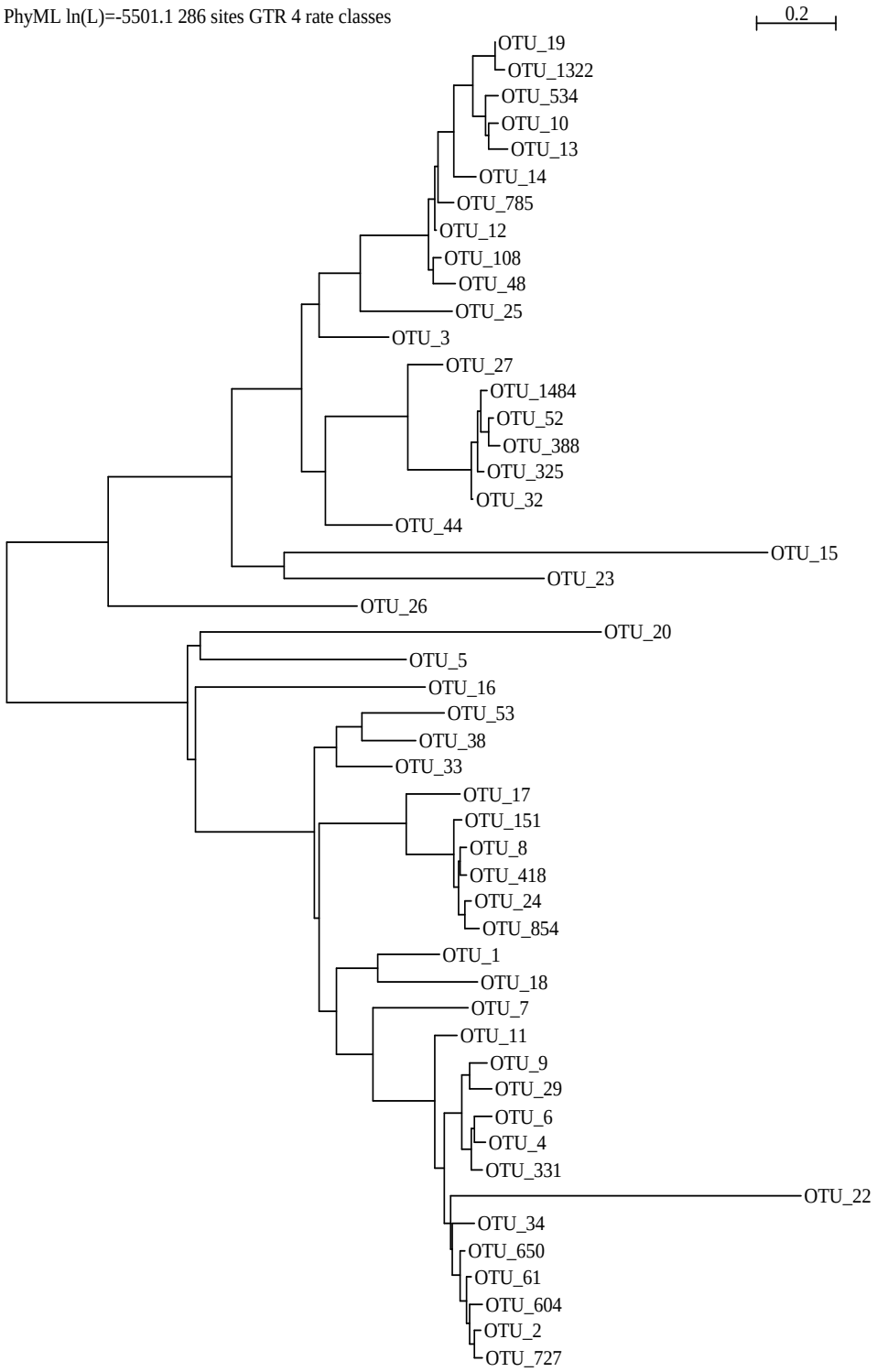


Figure 5a. Phylogenetic tree of aligned nucleotide sequences for the 50 most abundant *apr* OTUs.

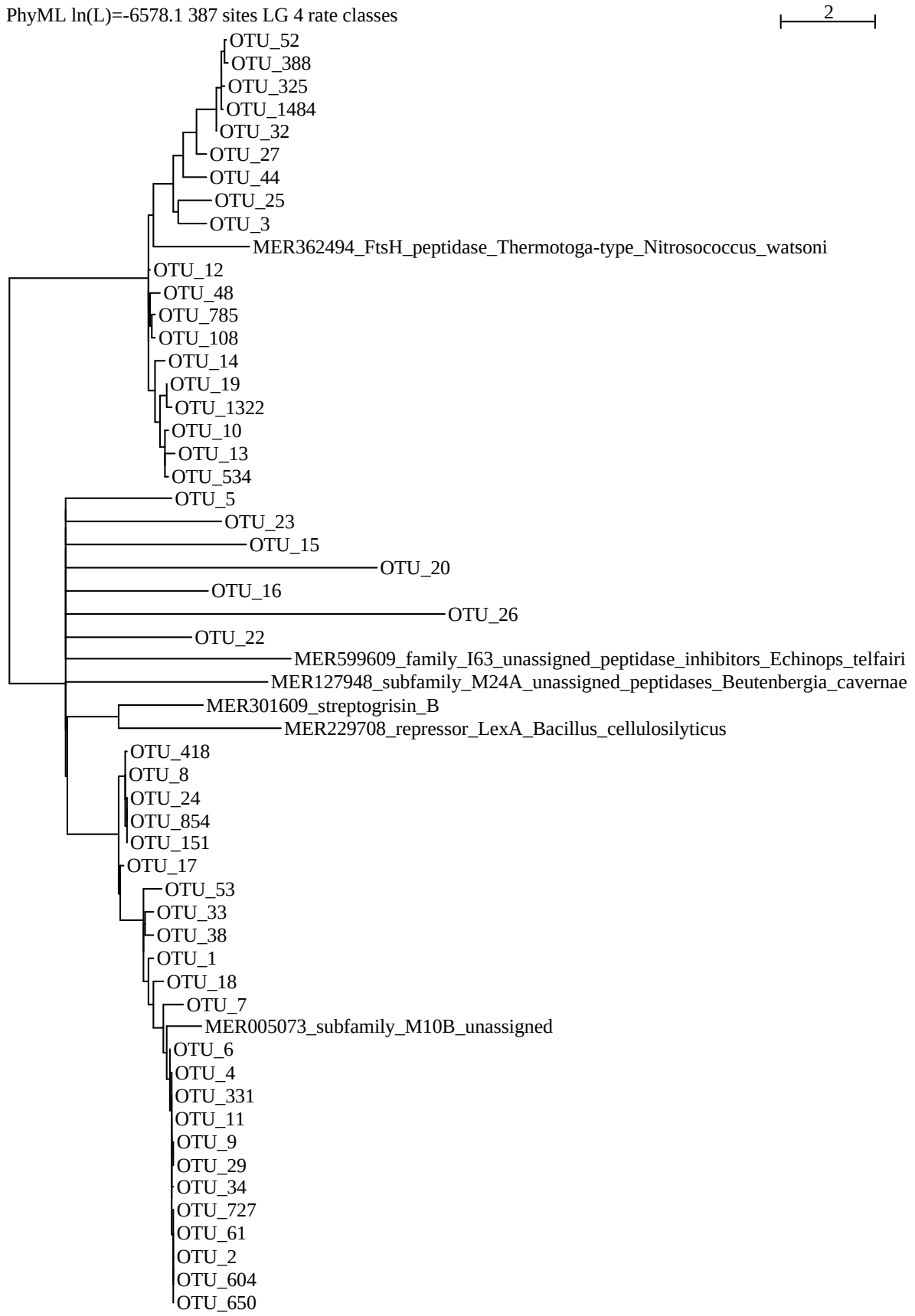


Figure 5b. Phylogenetic tree of aligned aminoacidic sequences for the 50 most abundant *apr* OTUs. Reference sequences of the most abundant blastx hits are also reported

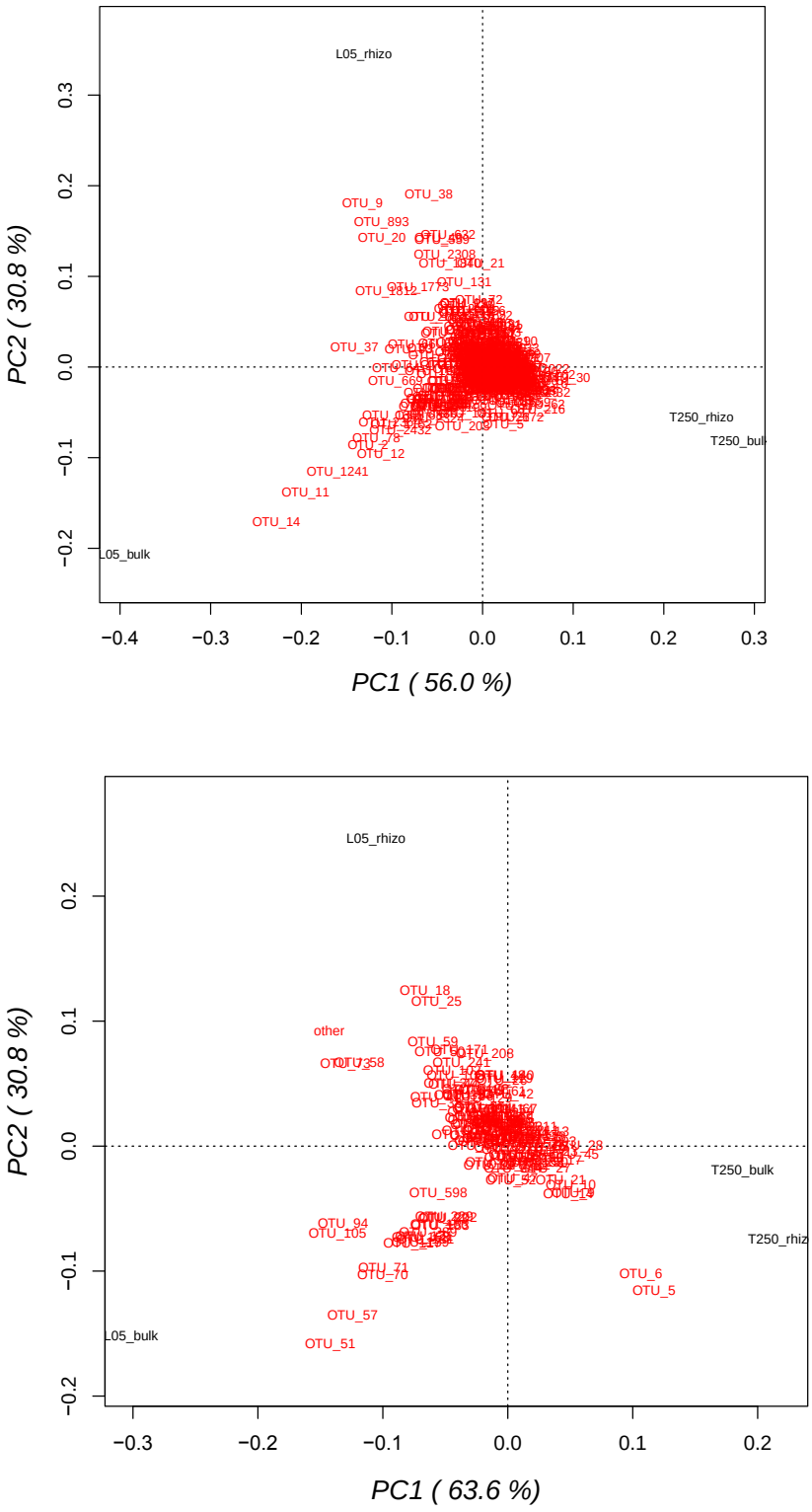


Figure 6 Principal Component Analysis of (a) *apr* OTU data and (b) npr OUT data covering the 99.9% of total diversity. Sample grouping is reported, together with ordiplot of OTUs scores.

Analysis of the most abundant *npr* OTUs revealed that majority of OTUs assigned to uncultured bacteria; most of the others showed high phylogenetic similarity with members *of Bacillus sp.* (Figure 4). Based on their sequences, the *npr* OTUs were more diverse than *apr* OTUs (Figure 4).

Multivariate analysis was conducted to explore the discrimination between samples, and to identify OTUs mostly responsible for differences. For the T250 variety, bulk and rhizosphere samples are closely grouped, while for L05 variety differences between rhizosphere and bulk soil *apr* OTUs patterns were higher. PCA also highlighted a number of OTUs that were more related to samples, especially for L05 bulk and rhizosphere (Figure 5).

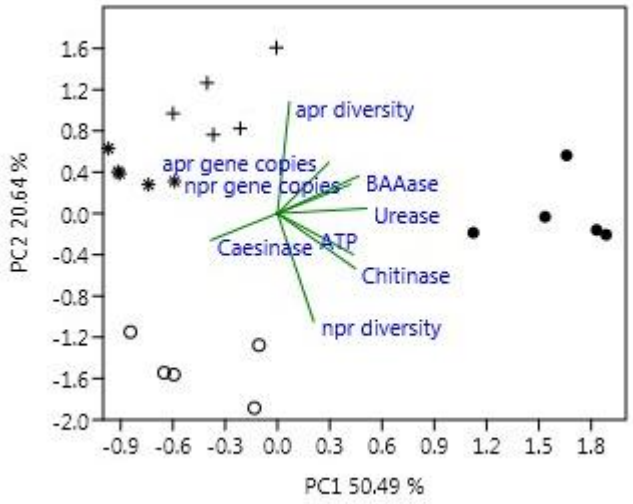


Figure 7. PCA on BAA-ase, Caesinase, Chitinase,Urease, ATP, npr gene copy numbers, *apr* gene copy numbers and diversity of *apr* and npr genes based on Shannon-Weiner indices of DGGE bands. Solid dots represent L05 rhizosphere samples, cross represent L05 bulk samples, circles represent T250 rhizosphere samples and stars represent T250 bulk samples

2.5 Discussion

With the exception of the casein hydrolyzing activity, all enzymatic activities and microbial biomass were found to be higher in the rhizosphere of the L05 maize as compared to T250 maize line, indicating that the high NUE L05maize line has a higher N mineralization rate in the rhizosphere than the low NUE T250 maize line. These results are in agreement with previous reports on the greater capability of the L05 maize line to acquire N from the rhizosphere (Zamboni et al., 2015; Pathan et al., 2015). The rhizosphere of the L05 and T250 maize lines also differed for the type of protease activities, as the L05 rhizosphere displayed a higher BAA- hydrolyzing activity whereas the rhizosphere of the T250 had a higher casein hydrolyzing activity (Figure 1). It is important to note that casein hydrolysing activities probably measured protease activity acting on high molecular weight substrates that generally are associated to microbial death events or release of extracellular enzymes degrading organic polymers (Nannipieri et al., 2012). These differences could depend on different factors including genetic diversity of the protease encoding genes, molecular integrity and extracellular stabilization of different proteases by the rhizosphere organic matter (Bonmati et al., 2009, Overall these results indicated that in the rhizosphere of the two maize lines the protein N mineralization depended on different proteolytic mechanisms.

Analysis of the DGGE fingerprints indicated higher complexity of the proteolytic communities in the rhizosphere of the L05 than those of the T 250 maize, showing that the two plant lines selected different proteolytic populations during the plant growth. These results are in line with those of Sakurai et al. (2007) who also reported rhizosphere effects on the diversity of the *apr* gene. Gene copy numbers were also significantly affected for the *apr* gene. These results support overall positive rhizosphere effect of high NUE on the *apr*

as compared to the *npr*, as shown by the significantly higher *apr* abundance in rhizosphere of L05 than T250 (Table 2). Previous studies on Maize rhizopshere by Aira et al. (2010), revealed that different genotypes modifies the structure of rhizospheric microbial communities, but not their abundance and no significant changes in biomass of main microbial groups were reported. But in our studies we have noticed significant changes in copy numbers of *apr* gene, but no significant changes in abundance of *npr* gene.

Our results based on the composition of the proteolytic community of the rhizosphere and bulk soil of the two maize lines indicate a significantly higher richness for *npr* than *apr* gene, and significant differences between rhizosphere of L05 and T250 maize lines. Analysis of OTUs confirmed results by Watanabe and Hayano (1994a, b) that *Bacillus* spp. are the main source of *npr* genes in soil. However, several unknown metallo-peptidase *npr* gene sequences outnumbered other known OTUs in both rhizosphere and bulk of the studied maize lines. This is indeed the first work dealing with the high-throughput assessment of protease genes in bulk and rhizosphere soils. Results indicate a high diversity of these genes in soil, as shown by the number of unique sequences and OTUs. However, together with the high number of unassigned sequences suggest that our current knowledge on the abundance and distribution of the protease encoding genes in soil is still very limited. Taken together, the genetic and biochemical analysis of the rhizosphere of the both maize lines indicated that the L05 maize line with higher NUE selected more strongly the proteolytic microbial communities in the rhizosphere as compared to the low NUE T250 maize line, with potential influence on the predominant protease mechanism. In fact, while the BAA hydrolyzing activity has a trypsin-like protease activity, the casein hydrolyzing activity is less specific serine proteases (Ladd, 1972). It can't be excluded that a more specialized

proteolytic community may contribute to the observed higher NUE of the L05that the T250 maize line.

For *apr*, the most abundant OTUs were reported from different members of *Pseudomonas sp.*; this confirms previous studies reporting high proteobacteria populations in maize rhizopshere (Peiffer et al., 2013). Furthermore these OTUs were significantly more abundant in the rhizosphere of the L05 than in the T250 rhizosphere. Other abundant *apr* OTUs detected in the maize rhizosphere such as *S*. *griseus* and *Caulobacter sp*., *N*. *watsoni* and *Clostridium* sp., *Brevibacillus* sp. and *Thermoactinomyces* sp. play important roles in maize growth, being involved in chitinase activity, plant pathogen biocontrol, non-symbiotic N fixationNO_3^--N reduction, or N and P mineralization (Jackson et al., 1997; Philippot et al. 2002; Bressan and Figueiredo, 2008; Peiffer et al., 2013; Yadav et al., 2013; Li et al., 2014). Interestingle another dominant OUT was *Dickeya* sp., a plant pathogen, also detected in maize rhizosphere (Chaparro et al., 2014); this may be related to the past use of the soil for maize cultivation.

Very interestingly many identified organisms contributing to both *apr* and npr OTUs, like Bacillus sp, Paenibacillus sp, Clostridium sp., Pseudomonas sp., Azoarcus sp., are plant growth promoting rhizobacteria (PGPR) (Hurek and Reinhold-Hurek, 2003, Kumar et al., 2011, Goswami et al., 2015, Kefela et al., 2015). Certain plant growth promoting microbes have been found to enhance N uptake from soil, primarily by nutrient mobilization and increase plant NUE (Parra-Cota et al., 2014). Present results also support the fact that most soil proteolytic communities play an important role as PGPRs, thus supporting their roles in soil fertility.

In conclusion, our work showed that maize line differing for NUE also host different microbial communities and select different protease encoding genes in their rhizosphere. In particular, the two maize lines mainly influenced the abundance and diversity of the *apr* gene than *npr* gene. Though npr gene is less affected by rhizosphere and plant properties, it has been unraveled that most *npr* OTUs were from unknown organisms and this suggests the need for a future research identifying hidden players behind npr gene pool. NUE dependent selective effect also results in differences in the functional potential of the rhizosphere microbial communities and apparently in the mechanisms responsible for the protein N mineralization. Future research should also characterize the N forms in the rhizosphere of the two maize lines and the maize root exudate profiles to further clarify the link between the protease gene diversity and the protein N fate in the rhizosphere of the studied maize lines.

References

Adamczyk, B., Smolander, A., Kitunen, V., Godlewski, M., 2010. Proteins as nitrogen source for plants. Plant Signaling and Behavior 5(7), 817–819.

Aira, M., Gómez-Brandón, M., Lazcano, C., Bååth, E., Domínguez , J., 2010. Plant genotype strongly modifies the structure and growth of maize rhizosphere microbial communities. Soil Biology and Biochemistry 42, 2276-2281.

Ascher, J., Ceccherini, M.T., Pantani, O.L., Agnelli, A., Borgogni, F., Guerri, G., Nannipieri, P., Pietramellara, G., 2009. Sequential extraction and Genetic fingerprinting of a forest soil metagenome. Applied Soil Ecology 42, 176-181.

Bach, H.J., Munch, J.C., 2000. Identification of bacterial sources of soil peptidases. Biology and Fertility of Soils 31(3); 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(3), 235-45.

Baudoin, E., Benizri, E., Guckert, A., 2002. Impact of growth stage on the bacterial community structure along maize roots, as determined by metabolic and genetic fingerprinting. Applied Soil Ecology19 , 135–145.

Baudoin, E., Benizri, E., Guckert ,A., 2003. Impact of artificial root exudates on the bacterial community structure in bulk soil and maize rhizosphere. Soil Biology and Biochemistry 35, 1183–1192.

Bonmatí, M., Ceccanti, B., Nannipieri, P., Valero, J., 2009. Characterization of humus–protease complexes extracted from soil. Soil Biology and Biochemistry 41(6), 1199–1209.

Bressan, W., Figueiredo, J.E.F., 2008. Efficacy and dose-response relationship in biocontrol of Fusarium disease in maize by Streptomyces spp. European Journal of Plant Pathology 120, 311-316.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., Knight R., 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of National Academy of Sciences of the United States of America 108, Suppl 1, 4516-22.

Chan,V., Dreolini, L.F., Flintoff, K.A., Lloyd, S.J., Mattenley, A.A., 2002. The Effect of Increasing Plasmid Size on Transformation Efficiency in Escherichia coli. Journal of Experimental Microbiology and Immunology l. 2, 207-223.

Chaparro, J.M., Badri, D.V., Vivanco, J.M., 2014. Rhizosphere microbiome assemblage is affected by plant development. The ISME Journal 8, 790-803.

Ciardi, C., Nannipieri, P., 1990. A comparison of methods for measuring ATP in soil. Soil Biology and Biochemistry 22, 725–727.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674-3676.

Dandeniya, W.S., Thies, J.E., 2012. Rhizosphere Nitrification and Nitrogen Nutrition of Rice Plants as affected by Water Management. Tropical Agricultural Research 24 (1), 1- 11.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460-2461.

Edgar, R.C., 2013.UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10, 996–998.

Fitz, W.J., Wenzel, W.W., Weishammer G., Blaž I., 2003. Microtome sectioning causes artifacts in rhizobox experiments. Plant and Soil 256, 455-462.

Fuka, M.M., Engel, M., Hagn, A., Munch, J.C., Sommer, M., Schloter, M., 2009. Changes of Diversity Pattern of Proteolytic Bacteria over Time and Space in an Agricultural Soil. Microbial Ecology 57,391-401.

Fuka, M.M., Engel, M., Gattinger, A., Bausenwein,U., Sommer, M., Munch, J.C., Schloter, M., 2008a. Factors influencing variability of proteolytic genes and activities in arable soils. Soil Biology and Biochemistry 40 (7), 1646–1653.

Fuka, M.M., Engel, M., Haesler, F., Welzl, G., Munch, J. C., Schloter, M., 2008b. Diversity of proteolytic community encoding for subtilisin in an arable field: spatial and temporal variability. Biology and Fertility of Soils 45, 185–191.

Godlewski, M., Adamczyk, B., 2007. The ability of plants to secrete proteases by roots. Plant Physiology and Biochemistry 45(9), 657–664.

Gooday, G.W., 1994. Physiology of microbial degradation of chitin and chitosan. Biodegradation 1, 177-19.

Goswami, D., Parmar, S., Vaghela, H., Dhandhukia, P., Thakker, J.N., 2015. Describing *Paenibacillus mucilaginosus* strain N3 as an efficient plant growth promoting rhizobacteria (PGPR). Cogent Food and Agriculture 1, 1000714.

Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. Molecular Biology and Evolution 27, 221-224

Guindon, S., Gascuel, O., 2003. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Systematic Biology 52, 696-704.

Hammer, Ø., Harper, D.A.T., Ryan, P.D., 2001. PAST: Paleontological statistic software package for education and data analysis. Palaeontologia Electronica 4,1–9.

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., 1993. Protease Activity in a Paddy Field Soil: Origin and Some Properties. Soil Science and Plant Nutrition 39 (3), 539-546.

Hirel, B., Le Gouis, J., Ney, B. Gallais, A., 2007. The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. Journal of Experimental Botany 58, 2369-2387.

Hurek, T., Reinhold-Hurek, B., 2003. *Azoarcus* sp. strain BH72 as a model for nitrogen-fixing grass endophytes. Journal of Biotechnology 106, 169–178.

Jan, M.T., Roberts, P., Tonheim, Jones, D.L., 2009. Protein breakdown represents a major bottleneck in nitrogen cycling in grassland soils. Soil Biology and Biochemistry 41, 2272-2282.

Kalisz, H.M., 1988. Microbial Proteinases. Advances in Biochemical Engineering/Biotechnology 36, 1–65.

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

Katharina, S., Stefan, H., Reingard, G., 2012. Direct cloning in Lactobacillus plantarum: Electroporation with non-methylated plasmid DNA enhances transformation efficiency and makes shuttle vectors obsolete. Microbial Cell Factories 11:141.

Kefela, T., Gachomo, E. W., Kotchoni, S. O., 2015. Paenibacillus polymyxa, Bacillus licheniformis and Bradyrhizobium japonicum IRAT FA3 Promote Faster Seed Germination Rate, Growth and Disease Resistance under Pathogenic Pressure. Journal of Plant Biochemistry and Physiology 3, 2329-9029.

Kumar, A., Prakash, A., Johri, B.N., 2011. Bacillus as PGPR in Crop Ecosystem. Chapter 2, Bacteria in Agrobiology. Crop Ecosystems XII, 37-59.

Jackson, A.M., Poulton P.R., Ball, S.A., 1997. Importance of farming practice on the isolation frequency of *Thermoactinomyces* species. Soil Biology and Biochemistry 29, 207-210.

Ladd, J.N., 1972. Properties of proteolytic enzymes extracted from soil. Soil Biology & Biochemistry 4, 227- 239.

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

Li, Y.L., Fan, X.R., Shen, Q.R., 2008. The relationship between rhizosphere nitrification and nitrogen-use efficiency in rice plants. Plant, Cell and Environment 3, 73-85.

Li, X., Rui, J., Xiong, J., Li, J., He, Z., Zhou, J., Yannarell, A.C., Mackie, R.I., 2014. Functional Potential of Soil Microbial Communities in the Maize Rhizosphere. PLoS One 9, e112609.

Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., Suzuki, A., 2010. Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. Annals of Botany 105,1141–1157.

McHugh, T.A., Schwartz, E., 2015. Changes in plant community composition and reduced precipitation have limited effects on the structure of soil bacterial and fungal communities present in a semiarid grassland. Plant and Soil 388,175-186.

Metcalfe, A.C., Krsek, M., Gooday, G.W., Prosser, J.I., Wellington, E.M.H., 2002. Molecular Analysis of a Bacterial Chitinolytic Community in an Upland Pasture. Applied And Environmental Microbiology 68 , 5042–5050

Mooshammer, M., Wanek, W., Hämmerle, I., Fuchslueger, L., Hofhansl, F., Knoltsch, A., Schnecker, J., Takriti, M., Watzka, M., Wild, B., Keiblinger, K.M., Zechmeister-Boltenstern, S.,

Richter, A., 2014. Adjustment of microbial nitrogen use efficiency to carbon:nitrogen imbalances regulates soil nitrogen cycling. Nature Communications 5, Article number 3694.

Nannipieri, P., Ceccanti, B., Cervelli, S., Sequi, P., 1974. Use of 01 Mpyrophosphate to extract urease from a podzol. Soil Biology and Biochemistry 6, 359–362.

Nannipieri, P., Ceccanti, B., Cervelli, S., Matarese, E., 1980. Extraction of phosphatase, urease, proteases, organic carbón, and nitrogen from soil. Soil Science Society of America Journal 44, 1011-1016.

Nannipieri, P., Paul, E., 2009.The chemical and functional characterization of soil N and its biotic components. Soil Biology and Biochemistry 41, 2357-2369

Nannipieri, P., Giagnoni, L., Renella, G., Puglisi, E., Ceccanti, B., Masciandaro, G., Fornasier, F., Moscatelli, M.C., Marinari, S., 2012. Soil enzymology : classical and molecular approaches. Biology and Fertility of Soils 48, 743-762.

Ngezimana, W., Agenbag, G.A., 2014. Effects of nitrogen and sulphur on seedling establishment, vegetative growth and nitrogen use efficiency of canola (Brassica napus L.) grown in the Western Cape Province of South Africa. Journal of Cereals and Oilseeds 5(2), 4-11.

Parra-Cota, F.I., Peňa- Cabriales, J.J., Santos-Villalobos, S., Martı´nez-Gallardo, N.A., De´lano-Frier, J.P., 2014. Burkholderia ambifaria and B. caribensis Promote Growth and Increase Yield in Grain Amaranth (Amaranthus cruentus and A. hypochondriacus) by Improving Plant Nitrogen Uptake. PLOS ONE 9, 1-14.

Pathan, S.I., Ceccherini, M.T., Pietramellara, G., Puschenreiter, M., Giagnoni, L., Arenella, M., Varanini, Z., Nannipieri, P., Renella, G., 2015. Enzyme activity and Microbial community structure in the rhizosphere of two maize lines differing in N use efficiency. Plant Soil 387, 413-424.

Peiffer, J.A., Spor, A., Koren, O., Jin, Z., Tringe, S.G., Dangl, J.L., Buckler, E.S., Ley, R.E., 2013. Diversity and heritability of the maize rhizosphere microbiome under field conditions. Proceedings of National Academy of Sciences of the United States of America 110, 6548-6553.

Pester, M., Rattei, T., Flechl, S., Gröngröft, A., Richter, A., Overmann, J., Reinhold-Hurek, B., Loy, A., Wagner, M., 2012. amoA-based consenus phylogeny of ammonia-oxidizing archaea and deep sequencing of amoA genes from soils of four different geographic areas. Environmental Microbiology 14, 523-539.

Philippot, L., Piutti, F., Martin-Laurent, F., Hallet, S., Germon, J.C., 2002. Molecular analysis of the nitrate-reducing community from unplanted and maize-planted soils. Applied and Environmental Microbiology 68, 6121-6128.

Rasche, F., Musyoki, M.K., Röhl, C., Muema, E.K., Vanlauwe, B., Cadisch, G., 2014. Lasting influence of biochemically contrasting organic inputs on abundance and community structure of total and proteolytic bacteria in tropical soils. Soil Biology and Biochemistry 74, 204-213.

Raun, W., Johnson, G.V., 1999. Improving nitrogen use efficiency for cereal production. Agronomy Journal 91, 357-363.

Raup, D.M., Crick, R.E., 1979. Measurment of faunal similarity in paleontology. Journal of Paleontology 53, 1213–1227.

Rawlings, N.D., Barrett, A.J., Bateman, A., 2012. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Research 40, D343-350.

Renella, G., Landi, L., Valori, F., Nannipieri, P., 2007. Microbial and hydrolase activity after release of low molecular weight organic compounds by a model root surface in a clayey and a sandy soil. Applied Soil Ecology 36, 124-129.

Sakurai, M., Suzuki, K., Onodera, M., Shinano, T.,′ Osaki, M., 2007. Analysis of bacterial communities in soil by PCR–DGGE targeting protease genes. Soil Biology and Biochemistry 39, 2777–2784.

Schimel, J.P., and Bennett , J., 2004. Nitrogen mineralization: challenges of a changing paradigm. Ecology 85(3), 591-602.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75, 7537-7541.

Shannon, C.E., Weaver, W., 1963. The mathematical theory of communication. University of Illinois Press, Urbana, IL.

Sun, J., Zhang, Q., Zhou, J., Wei, Q., 2014. Illumina amplicon sequencing of 16S rRNA tags reveals bacterial community development in the rhizosphere of apple nurseries at a replant disease site and a new planting site. PLoS ONE 9(10), e111744.

Tsuboi, S., Yamamura, S., Imai, A., Satou, T., Iwasaki, K., 2014. Linking Temporal Changes in Bacterial Community Structures with the Detection and Phylogenetic Analysis of Neutral Metalloprotease Genes in the Sediments of a Hypereutrophic Lake. Microbes and Environments 29, 314-321.

Vasileiadis, S., Puglisi, E., Arena, M., Cappa, F., van Veen, J.A., Cocconcelli, P.S., Trevisan, M., 2013.Soil microbial diversity patterns of a lowland spring environment. FEMS Microbiology Ecology 86:172-184.

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

Watanabe, K., Hayano, K., 1994b. Source of Soil Protease Based on the Splitting Sites of a Polypeptide. Soil Science and Plant Nutrition 40 (4), 697-701.

Watanabe., K., 2009. Detection of protease genes in field soil applied with liquid livestock feces and speculation on their function and origin. Soil Science and Plant Nutrition 55, 42–52.

Weintraub, M.N., Schemel, J.P., 2005. Seasonal protein dynamics in Alaskan arctic tundra soils. Soil Biology and Biochemistry 37, 1469-1475.

Wenzel, W.W., Weishammer, G., Fitz, W.J. , and Puschenreiter, M., 2001. Novel rhizobox design to assess rhizosphere characteristics at high spatial resolution. Plant and Soil 237, 37-45.

World reference base for soil resources 2006 - A framework for international classification, correlation and communication , Rome, Food and Agriculture Organization of the United Nations; 10013/epic.43321

Xu, G., Fan, X., Miller, A.J. , 2012. Plant Nitrogen Assimilation and Use Efficiency. Annual reviews of Plant Biology 63, 153-182.

Yadav, H., Gothwal, R.K., Nigam, V.K., Sinha-Roy, S., Ghosh, P., 2013. Optimization of culture conditions for phosphate solubilization by a thermo-tolerant phosphate-solubilizing bacterium *Brevibacillus* sp. BISR-HY65 isolated from phosphate mines. Biocatalysis and Agricultural Biotechnology 2, 217-225.

Zamboni, A., Astolfi, S., Zuchi, S., Pii, Y., Guardini, K., Tononi, P., Varanini, Z., 2014. Nitrate induction triggers different transcriptional changes in a high and a low nitrogen use efficiency maize inbred lines. Journal of Integrative Plant Biology 56: 1080–1094.

Zhang, F., Shen, J., Li, L., Liu, X., 2004. An overview of rhizosphere processes related with plant nutrition in major cropping systems in China. Plant and Soil 260, 89–99.

3. Soil microbial diversity, protease encoding genes and ammonium monooxygenase genes in response to elevated atmospheric Carbon dioxide

3.1 Abstract

Elevated CO_2 (eCO_2) results in N-limitation that brings about increase in N-immobilization and a reduced N-mineralization. Proteases are main enzymes responsible for N-mineralization and previous studies suggest close association of reduced N-mineralization with shifts in ammonium oxidation pathways under elevated CO_2. We have hypothesized that the microbial communities respond by changing their community structure under N-limitation resulting from high atmospheric CO_2, it should also affect the diversity and abundance of genes involved in proteolysis and in ammonium oxidation. This study was carried out to understand the effect of elevated CO_2 on the microbial communities, proteolytic genes and bacterial amoA genes mediated via N-limitation in the eCO_2 system. A Free Air Carbon dioxide Enrichment (FACE) field was established and maintained at 550 ppm CO_2 and in parallel a control plot having ambient atmospheric CO_2 was maintained. On both the plots Cappelli cultivar of wheat was grown. Rhizosphere and bulk soil from FACE field and from control field were sampled to proceed with further studies. To study the diversity of bacterial alkaline metallo-peptidase genes (*apr*) and bacterial ammonium monooxygenase genes (amoA), we have used the PCR-DGGE approach. Gene abundance was studied using quantitative PCR (qPCR). To look for shifts in microbial communities we have applied Illumina sequencing of 16S genes of bacteria. DGGE results show a shift in diversity of *apr* genes and amoA genes. Our results shows a significant ($P<0.05$) reduction in gene copy numbers of both protease genes and of amoA genes. Reduced abundance was noticed not only in rhizosphere soil but also in bulk soil. Illumina sequencing results also

shows a shift in communities. Results suggest that the proteolytic communities in soil adapt themselves in response to rising CO_2 in soil, particularly in rhizosphere.

Keywords: elevated CO2, FACE, protease encoding gene, qPCR, PCR-DGGE,

3.2 Introduction

Microbial communities are important drivers of biogeochemical cycles in soil and play an important role in soil functioning. For ecosystem processes, functional diversity is a more common measure than taxonomic diversity and functional diversity can also be measured by measuring functional genes that play roles in ecosystem processes. Structure of microbial communities and their functional diversity is affected by different C and N inputs to the soil (Minz et al 2013). Increased atmospheric CO_2 leads to increased C input in soil (Kessel et al 2000, Adair et al. 2000, Jastrow et al. 2005) and also a shift in microbial community structure in soil (Ginkel et al. 2000). Increased atmospheric CO_2 not only changes C input but also affect the N availability and leads to N limitation in eCO_2 environments. Changes in N and C dynamics and an increased N immobilization and reduced N mineralization results in N limitation in eCO_2 exposed ecosystems (Schlesinger et al. 2006, Finzi et al., 2006). Protease activity is one of the major activities that bring about N mineralization by depolymerization of proteins and peptides (Nannipieri and Eldor, 2009). Most previous studies focused on protease enzyme activity under eCO_2 and depicted an increased proteolytic activity in soil under N limitation (Sims and Wander 2002, Kandler et al., 2006, Xuexia et al., 2006, Drissner et al., 2007), but there is a lack of study on genes encoding proteases under eCO_2 environment. Study of protease encoding genes should give information on real proteolytic potentials of soil under N limitation resulting from eCO_2 environment.

In soil concentrations of C and N regulate protease activity, an increased input of Carbon not only reduces protease activity but also induce shifts in NH_4^+ assimilation pathway (Geisseler and Howarth, 2008). Thus we expect that the genes involved in NH_4^+ metabolism should also respond to increased C input in soil. Consequently, the underlying hypothesis is that a reduced content of proteins in organic residues derived from crops cultivated under elevated CO_2 is inhibiting the abundance of proteolytic genes and thus subsequently depressing the abundance of genes involved in nitrification and denitrification. At molecular level not much information is available about the response in the structure and abundance for the proteolytic genes and ammonium monooxygenase (amoA) genes under eCO_2, to bridge this gap present studies were carried out. We have selected alkaline metallo-peptidase gene (*apr*) and amoA gene for this investigation.

Free-Air Carbon dioxide Enrichment (FACE) allows the fumigation of natural ecosystem and helps in understanding the changes in exposed ecosystem due to elevated CO_2 (eCO_2) (Allen et al. 1992, Ashenden et al. 1992). This approach allows study of a wide range of processes in their natural environment unlike other microcosm approaches, where natural conditions are manipulated. We used a FACE approach to study the changes in bacterial community structure and proteolytic gene diversity and abundance in the rhizosphere soil of Cappelli wheat cultivar, with the bulk soil from a plot exposed to 550 ppm CO_2. In parallel, soils from a plot with same conditions that of FACE plot, but with ambient atmospheric CO_2 (aCO_2) was studied as a control. Microbial community structure was studied using illumina sequencing of 16S amplicons and changes in proteolytic gene abundances were studied using qPCR. To study structure of proteolytic communities DGGE approach was applied.

3.3 Materials and methods

Study site and plot setup: Study site is located in Fiorenzuola d'Arda (44.927°N, 9.893°E). Soil was a fine silty,mixed, mesic Udic Ustochrepts with pH PH 7.9., 1.5 % total N,and 2.2 % organic matter. Wheat genotype Capelli was grown within the FACE facility of the Genomics Research Centre of the Consiglio per la Ricerca e sperimentazione in Agricoltura (CRA-GPG) at Fiorenzuola d'Arda (44.927°N, 9.893°E) applying a split plot design with FACE and control octagons distributed at random within the experimental field (3 FACE, 3 controls). The single FACE and control systems contained two blocks (northern and southern side) with plots (1.32 x 2.2 m) for the genotype as split plots. Cappelli is a variety with a prominent role in Italian durum wheat breeding. Sowing at optimal sowing time (October 19th 2011) was assured by a pre-harrowing irrigation due to dry soil conditions. The CO_2 mixing ratio for the FACE treatment target was fixed at 570 ppm representing a value within the upper range of scenarios for the mid Century atmospheric mixing ratio. FACE treatment was started on November 16th, 2011 and stopped when leaves were senescent at June 14th, 2012.

FACE treatment was interrupted for 20 days in February 2012 when the plots were covered with snow. Apart from the CO_2 fumigation, the experiment was performed according to standard local agronomic practice and with the objective to avoid major pests and diseases. The plots were fertilised with application of an N:P:K fertiliser at pre seeding and two top dressings with ammonium nitrate for a total of 149 kg N ha-1. Final harvest was carried out manually in July 2012.

Sampling: Rhizosphere and bulk samples were collected in triplicates from each, FACE plot and control plot, ending up in 12 samples. Immediately after collection samples were kept at 4°C, till they were transported to laboratory. In laboratory soil was sieved through a 2mm sieve and was preserved at -20°C till nucleic acid extraction.

For extraction 0.5 g soil was weighed in lysing matrix tube (MP Biomedicals) and sequential extraction of DNA was carried out as mentioned in Ascher et al 2011. Only intracellular fraction of DNA was used for further studies. DNA was kept at -20°C till further analyses.

Nucleic acids extraction and PCR-DGGE analysis: DNA was extracted by sequential extraction method from 0.5 g soil as described by Ascher et al. (2009) using the FastDNA spin kit for soil (MP Biomedicals, USA), and the intracellular DNA fraction was used in this study. The DNA yield and purity were analysed with a Qubit 2.0 fluorometer (Life Technologies, USA) using Quant-iT dsDNA *HS* kit according to the manufacturer's instructions, and stored at -20°C till analysis.

PCR-DGGE: The DGGE fingerprints were performed using an INGENY PhorU System (Ingeny International BV, Netherlands). Primer pair FP*apr*I/RP*apr*II for alkaline protease (*apr*) genes were adapted from Bach et al. 2001. The DGGE conditions for the fingerprinting of the *apr* amplicons were those previously used by Sakurai et al. (2007). For amoA gene primer pair were used as mentioned in Rotthauwe et al 1997 and the DGGE was performed according to Ceccherini et al 2007. The DGGE gels were stained with SybrGreen I (FMC Bio Products, Rockland, ME, USA), and the banding patterns were analysed by a Gel Doc system (Bio-Rad, USA).

Quantification of genes: Quantitative PCR (qPCR) was conducted on a CFX Connect Real-Time PCR Detection System (Bio-rad Laboratories) to determine the abundance of *apr* and genes. Each qPCR assay was conducted in a 96-well plate and included three replicates for each standard, negative controls, and sample. Amplification was performed using the iTaq

Universal SYBR Green Supermix (Bio-rad Laboratories), adding to each reaction mixture forward and reverse primers for both genes at concentration of 0.6 µM, 3% of bovine serum albumin (BSA), 20 ng DNA template for both genes. For *apr* gene primer sets FP *apr*I/ RP *apr*II were used according to Bach et al. (2001). *Pseudomonas fluorescence* (isolated from an agricultural soil) was used as the source of positive control for *apr* gene. The PCR runs for *apr* gene started with an enzyme activation step at 95°C for 3 min, followed by 42 cycles of denaturation at 94°C for 25 s., annealing was carried out at 54°C for 30 s, followed by extension at 72°C for 30 s. For amoA gene primer pair amoA1F and amoA2R were used according to Rotthauwe et a., 1997 and N. multiformis (ATCC 25196) was used as positive control. The PCR runs for amoA gene also started with an enzyme activation step at 95°C for 3 min, followed by 42 cycles of denaturation at 94°C for 25 s., annealing was performed at 60°C for 30 s and an extension step at 72°C for 25s. The specificity of amplification products were confirmed by melting curve analysis and expected sizes of amplified fragments were checked by running the amplicons on a 2% agarose gel stained with ethidium bromide for 90 mins at 100 V.

16 S gene Sequencing: PCR for preparing amplicons for sequencing were prepared in two steps as mentioned in Berry et al 2011. Amplification was performed using Phusion Flash PCR Master Mix (Thermo Fischer Scientific), adding to each reaction mixture forward and reverse primers for both genes at concentration of 0.6 µM, 1 ng DNA template in a 25 µl reaction. Original DNA was diluted to 0.1ng/µl concentration to be used as template and 1 µl of template was used for PCR for bacterial 16 S genes. Primers used here target V3-V4 regions of bacterial 16 S sequence. 1^{st} step of PCR consisted of 20 cycles and for 2^{nd} step of PCR, which had 10 cycles, 1 µl amplicon from 1^{st} step were used as template. PCR mix was

prepared using reverse primer, Flash master mix and nuclease free water. Forward primer with barcode for each sample was added directly in PCR tube for each sample, 48 different barcodes were used that were unique to each sample. After 2nd step of PCR, DNA quantification was carried out using Qubit ds DNA high sensitivity method. Amplicons with barcodes resulted from 2nd step of sequencing having V3-V4 regions were sequenced on an Illumina HiSeq sequencer.

Reads were filtered using the RDP sequencing pipeline (Cole et al. 2009).Clustering, alignment and chimera removal were performed using QIIME software package (Caporaso et al. 2010). Each trimmed FASTA sequence was BLASTed against SILVA 16S database (Quast et al. 2013) and later analysis was performed on MOTHUR (Schloss 2009). Principal component analysis (PCA), statistical calculations and graphs were plotted on R (http://www.R-project.org.).

Statistical analyses: qPCR data were analyzed by using Biorad software and later ANOVA was performed and the significance of differences between mean values were determined by the Fisher PLSD. For PCR-DGGE analysis, bands were identified and their intensities were measured after normalizing lanes and background subtraction using Quantity-One® software (Bio-Rad Laboratories, USA). The DGGE banding pattern was clustered to UPGAMA dendrograms based on Raup and Crick similarity indices (Raup and Crick, 1979) using the Quantity-One® software. Range weighted richness (Rr), for DGGE bands were measured as mentioned in mazorati et al. 2008.

3.4 Results

PCR-DGGE

Range weighted richness (Rr) for DGGE bands was higher for bulk samples than rhizosphere samples (Figure 1a and 1b) for both the genes indicating rhizosphere selection for bands whereas bulk soil had less selected communities and more diversity.

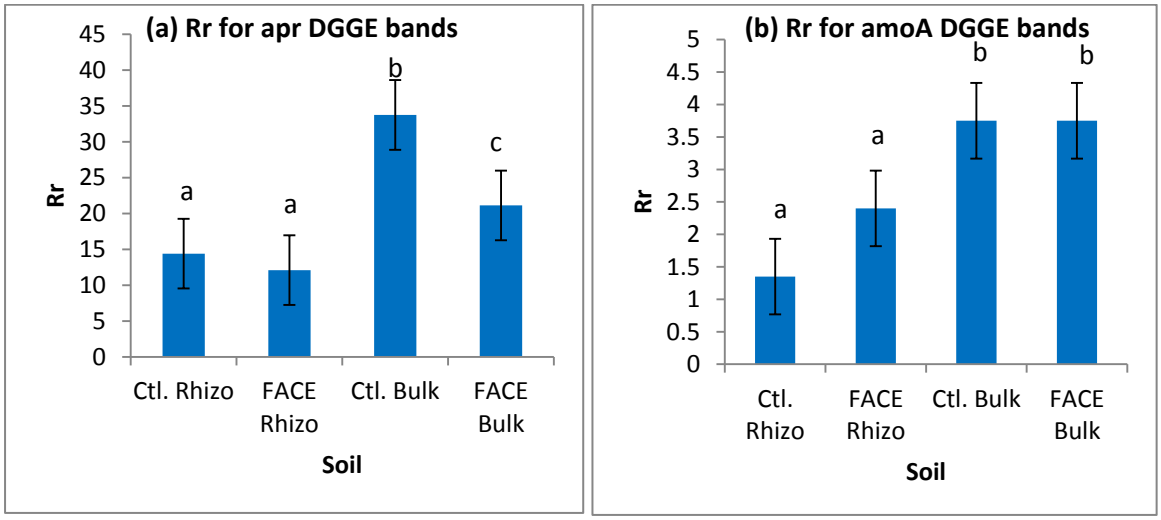
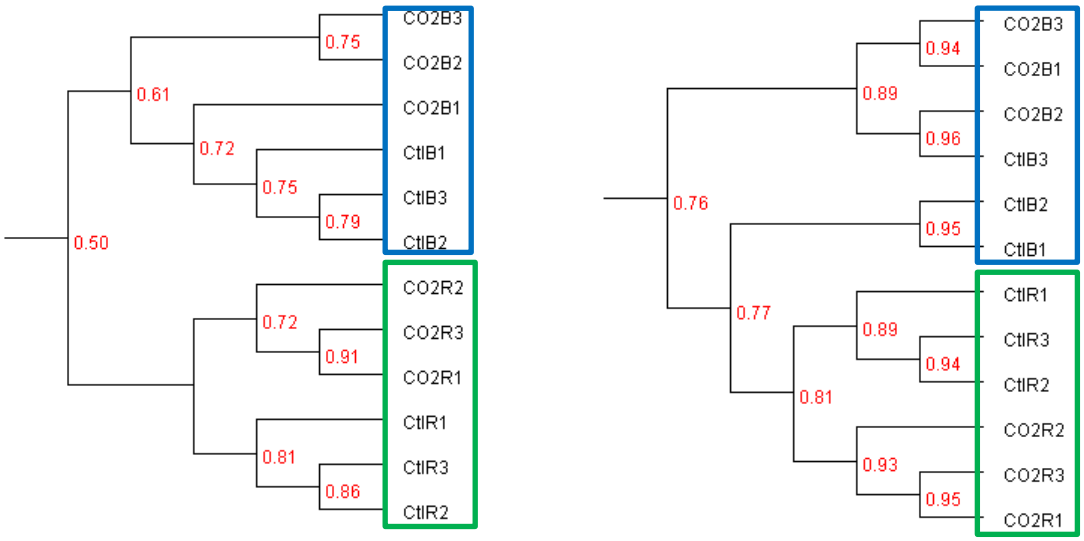


Figure 1. Range weighted richness (Rr) for DGGE bands of (a) *apr* gene and (b) amoA AOB gene. Standard deviations are shown by error bars and different superscripts indicate significant differences ($P < 0.05$) of values within each soil type.

UPGAMA cluster analysis for *apr* gene (Figure 1a) and amoA gene (Figure 1b) shows that irrespective of the treatment all the Rhizosphere samples clustered together whereas all the bulk samples clustered together. It indicates that the difference in diversity between rhizosphere and bulk soil were stronger that the difference in diversity between treatment or it could also mean that the effect of CO_2 treatment on these 2 genes was not strong enough to be identified using DGGE.

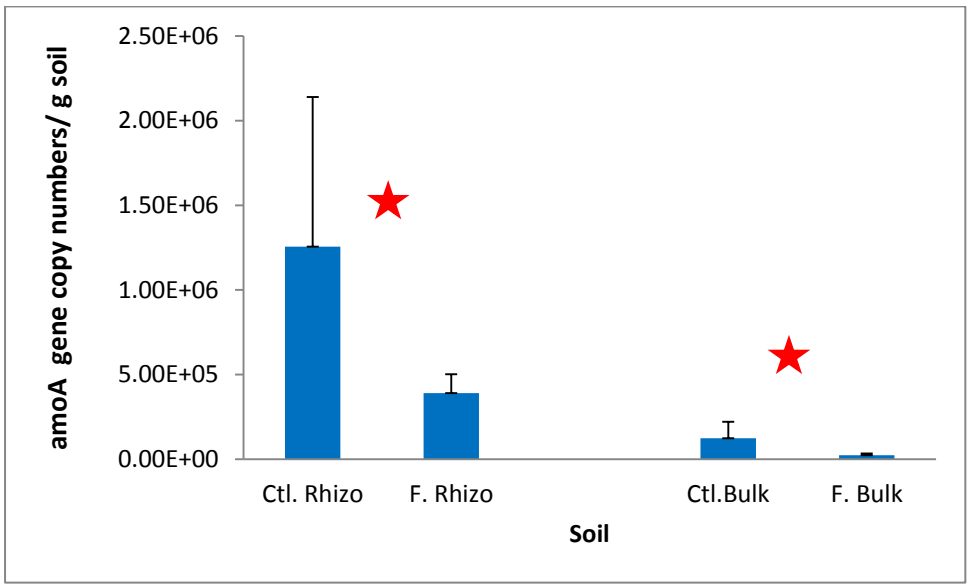


(a) (b)

Figure 2. UPGAMA clustering based on Raup Crick similarity of (a) *apr* gene and (b) amoA AOB gene.

qPCR

Both *apr* genes and amoA genes have shown a significant reduction in abundance ($P<0.05$) in FACE soil. This decrease in gene copy numbers is observed both in rhizosphere and in bulk soil (Figure 1a and 1b). Higher gene abundance was noticed for both *apr* and amoA genes in respective rhizosphere soils.



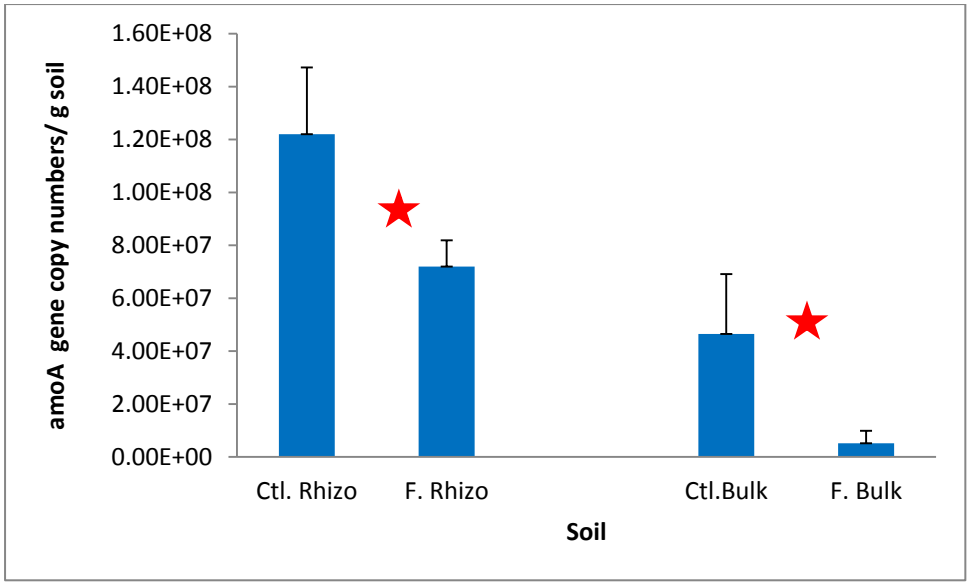


Figure 2: (a) *apr* gene copy numbers per gram of soil (b) amoA gene copy numbers per gram of soil. Standard deviations are shown by error bars and asterisk sign indicate significant differences ($P < 0.05$) of values within each soil type.

Sequence analysis results

Number of sequences retrieved for eCO_2 treated rhizosphere soil showed a significant ($p < 0.05$) reduction whereas for bulk soil number of sequences were not affected by the treatment with eCO_2 (Figure 3).

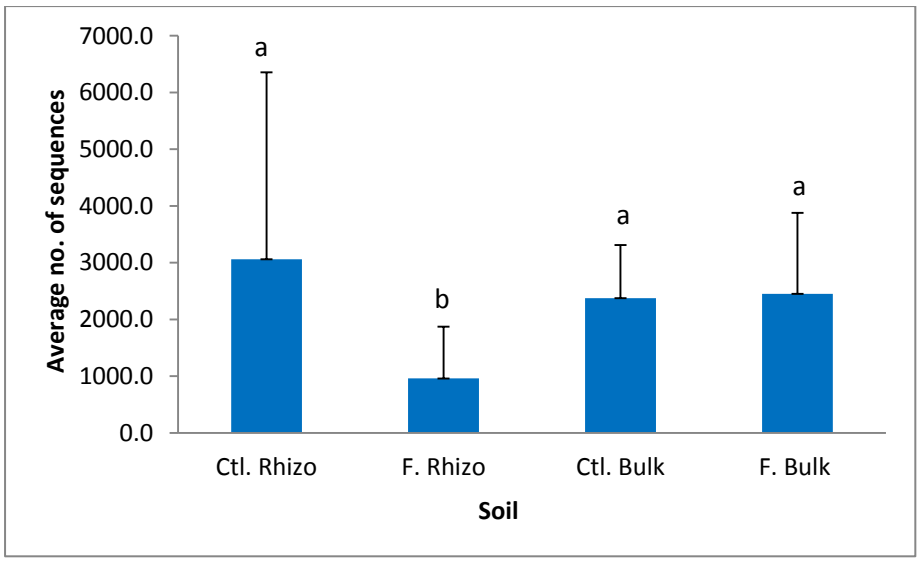


Figure 3. Average no. of sequences retrieved from each soil type

Standard deviations are shown by error bars and different superscripts indicate significant differences ($P < 0.05$) of values within each soil type.

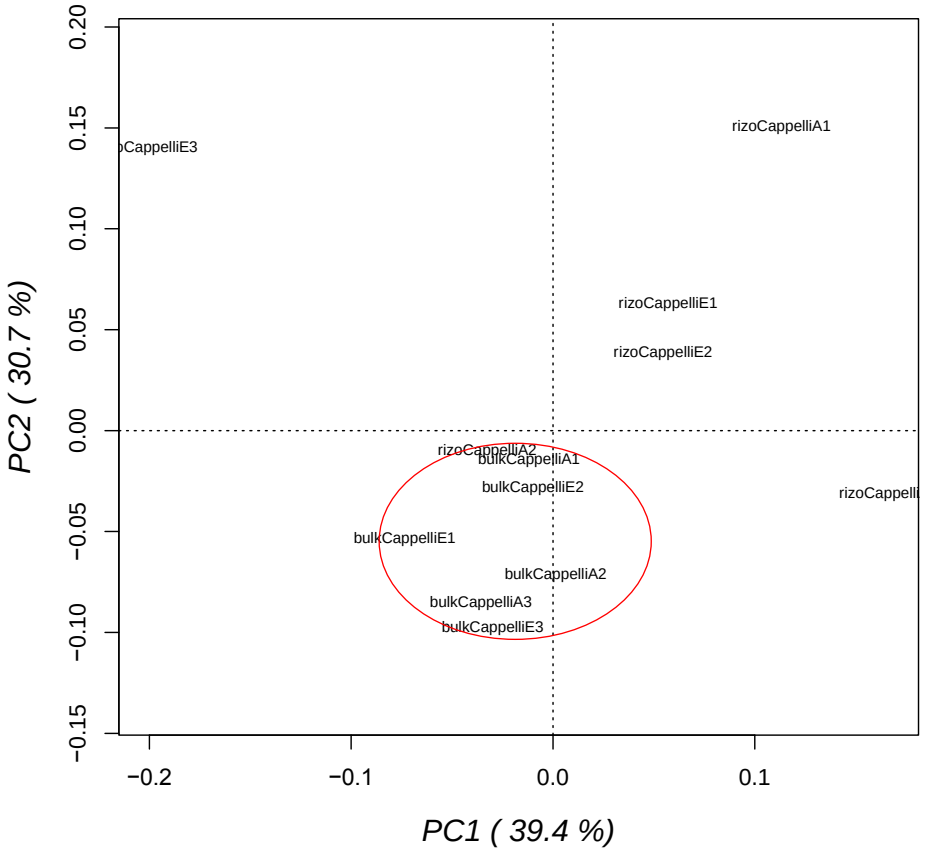


Figure 5. PCA on taxonomy data

Principal Component Analysis (PCA) on the taxonomic data showed that all the bulk samples irrespective of the treatment were clustering together whereas rhizosphere samples were discretely distriputed on a PCA plot (Figure 4), indicating that the bulk samples had very similar taxonomic makeup but the rhizosphere samples had complex heterogenous taxonomic diversity. There was a marked reduction in OTU diversity in eCO_2 treated rhizosphere samples as indicated by the Simpson's index of diversity values for different samples (Figure 6).

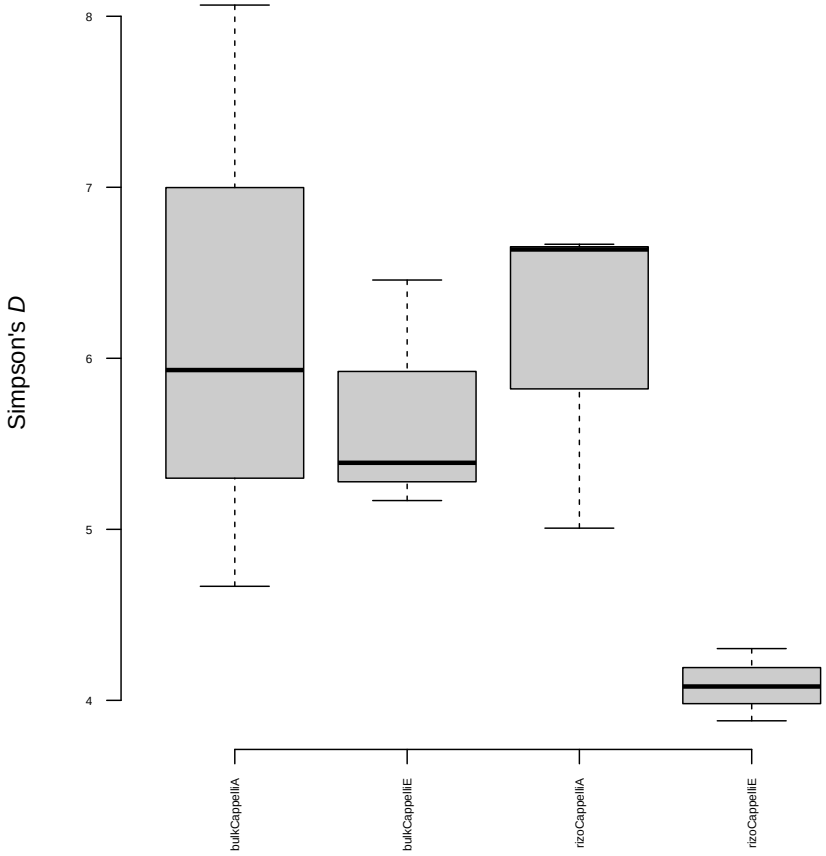


Figure 6. OTU-based analyses: Simpson index of diversity

3.5 Discussion

Results based on DGGE and qPCR for both bacterial genes (*apr* and amoA), indicate a higher diversity and abundance in rhizosphere soil than in bulk for both the eCO_2 treated and in aCO_2 envirnments. This study shows that irrespective of the treatment, rhizosphere soil harbours higher genetic diversity and abundance for bacterial genes involved in protease activity and for ammonium oxidation. This fact is supported by the previous findings that rhizo-deposition support higher microbial diversity in rhizosphere soil than in bulk (Morgan et al. 2005) and we in our studies noticed that this capacity of rhizosphere to harbour higher microbial populations in rhizosphere soil is not affected by the effect of eCO_2. We noticed a drop in abundance of genes in bulk soil too which indicates that the effect of eCO_2 can show its effect independent of plant mediated influence on soil. Plant root zone is a complex system, where interaction of microbes and plant is controlled by a plethora of different inputs from both plant and soil (Huang et al. 2014).

In our studies we noticed a reduction in the abundance of *apr* gene, which is possibly just one side of the story and there may be some other protease encoding gene taking over *apr*. Most previous studies noticed an increase in protease enzyme activity in soil (Sims and Wander 2002, Kandeler et al 2006). In our studies we have studied genes for only one of the many types of proteases. A drop in abundance of *apr* gene does not represent reduction in abundance of whole proteolytic gene communities. It possibly could be due to some other proteolytic gene take over the communities responsible for *apr* gene. We also observed a decrease in abundance of bacterial amoA genes on exposure to eCO2. Previous studies reported a reduction in amoA AOB genes (Kelly et al. 2013, Zhang et al. 2013), an increase

(He et al. 2014) or no response at all (Nelson et al. 2010), depending on other factors like N fertilization, crop type and soil.

Irrespective of the treatment, all the bulk samples clustered together in PCA graph for taxonomy, whereas rhizosphere samples were more discrete and scattered on PCA bi-plot. Exudates from roots affect the microbial community structure in the rhizosphere soil (Bais et al., 2006, Huang et al. 2014) and it is evident that microbial community structure and functions in rhizosphere are different from that of the bulk soil (Minz et al. 2013, Philipot et al, 2013). Results shows that rhizosphere samples had more diversity deviations within replicates, whereas bulk samples share similarity with each other. Simpson's index of diversity significantly reduced ($P<0.05$), in rhizosphere samples and this change in microbial diversity was expected as eCO_2 alters the microbial community composition in FACE field (Lesaulnier et al 2007) and the microbially mediated C and N cycling (Xu et al. 2013). Studies from He et al., 2012 reported a mixed response from different members of same phyla and class under the eCO2 effect. Some members of phyla responded by showing an increased abundance whereas some others from same phyla showed a reduction in diversity and abundance. Similar results were obtained for gamma-Proteobacteria which consist most members with *apr* gene like Pseudomonas spp. and Proteus mirabilis.

As gene abundance alone doesn't account for the expressed activity, an in-depth study of gene expression behaviour in high CO_2 should give some insights. Furthermore, the type of protease genes should also be studied to get a complete picture. Future research should focus also on the expression of the protease gene and a metatranscriptome study should reveal more about the shifts in overall pathways due to N-limitation.

References

Adair E.C., Reich P.B. , Trost J.J., Hobb S.E. (2011). Elevated CO2 stimulates grassland soil respiration by increasing carbon inputs rather than by enhancing soil moisture, Global change Biology, doi: 10.1111/j.1365-2486.2011.02484.x.

Allen, L.H., Drake, B.G., Rogers, H.H., Shinn J.H. (1992). Field techniques for exposure of plants and ecosystems to elevated CO2 and other trace gases, Critical reviews in plant science, 11 (2,3): 85-119.

Ascher J., Ceccherini M.T., Pantani O.L. , Agnelli A., Borgogni F., Guerri G., Nannipieri P., Pietramellara G. (2009) Sequential extraction and Genetic fingerprinting of a forest soil metagenome, Applied Soil Ecology 42, 176-181.

Ashenden T.W., Baxter R., Rafarel C. R. (1992). An inexpensive system for exposing plants in the field to elevated concentrations of CO_2, Plant, cell and environment, Volume 15, Issue 3, *Apr*il 1992, Pages 365–372.

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.

Bais H.P., Weir T.L., Perry L.G., Gilroy S., Vivanco J.M. (2006). The role of root exudates in rhizosphere interactions with plants and other organisms, Annual Review of Plant Biology, Vol. 57: 233-266.

Berry D, Mahfoudh KB, Wagner M and Loy A (2011). Barcoded primers used in multiplex amplicon pyrosequencing vias amplification. 77: 7846-7849.

Caporaso J.G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.D., Costello E.K., Fierer N., Pena A.G., Goodrich J.K., Gordon J.I., Huttley G.A., Kelley S.T., Knights D., Koenig J.E., Ley R.E., Lozupone C.A., McDonald D., Muegge B.D., Pirrung M., Reeder J., Sevinsky J.R., Turnbaugh P.J., Walters W.A., Widmann J., Yatsunenko T., Zaneveld J., Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7:335–336.
http://dx.doi.org/10.1038/nmeth.f.303.

Ceccherini M.T., Ascher J., Pietramellara G., Mocali S., Viti C., Nannipieri P. (2007). The effect of pharmaceutical waste-fungal biomass, treated to degrade DNA, on the composition of eubacterial and ammonia oxidizing populations of soil, Biology and fertility of soil 44; 299-306.

Cole J.R., Wang Q., Cardenas E., Fish J., Chai B., Farris R.J., Kulam-SyedMohideen A.S., McGarrell D.M., Marsh T., Garrity G.M., Tiedje J.M. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37:D141–D145. http://dx.doi.org/10.1093/nar/gkn879.

Drissner D., Blum H., Tscherko D., Kandeler E. (2007). Nine years of enriched CO2 changes the function and structural diversity of soil microorganisms in a grassland, European Journal of Soil Science, February 2007, 58

Evans, L. S. and Hendrey, G. R. (1992). Responses of Cotton Foliage to Short-Term Fluctuations in Co-2 Partial Pressures. Critical Reviews in Plant Sciences 2-3, 11: 203-212.

Finzi A.C., Moore D.J. P., DeLucia E.H., Lichter J., Hofmockel K.S., Jackson R.B., Kim H.S., Matamala R., McCarthy H.R., Oren R., Pippen J.S., Schlesinger W.H. 2006. Progressive Nitrogen Limitation of Ecosystem Processes under elevated CO_2 in a warm-Temperate Forest. Ecology 87:15–25.http://dx.doi.org/10.1890/04-1748.

Geisseler D., Horwath W.R. (2008). Regulation of extracellular protease activity in soil in response to different sources and concentrations of nitrogen and carbon, Soil Biology and Biochemistry, Volume 40, Issue 12, December 2008, Pages 3040–3048.

Ginkel J.H.V., Gorissen A., Polci D. (2000). Elevated atmospheric carbon dioxide concentration: effects of increased carbon input in a *Lolium perenne* soil on microorganisms and decomposition, Soil Biology and Biochemistry 32(4); 449-456.

He Z. , Xiong J., Kent A.D. , Deng Y., Xue K., Wang G., Wu L., Nostrand J.D. V., Zhou J. (2014). Distinct responses of soil microbial communities to elevated CO_2 and O_3 in a soybean agro-ecosystem, *The ISME Journal* 8, 714–726.

Hendrey, G. R. (1992). The Doe/Usda Face Program - Goal, Objectives, and Results through 1989. Critical Reviews in Plant Sciences 2-3, 11: 75-83.

Huang X.F., Chaparro J.M., Reardon K.F., Zhang R., Shen Q., Vivanco J.M. (2014). Rhizosphere interactions: root exudates, microbes, and microbial communities, Botany 92: 267–275.

Jastrow J.D., Michaelmiller R., Matamala R., Norby R.J., Boutton T.W., Rice C.W., Owensby C.E. (2005). Elevated atmospheric carbon dioxide increases soil carbon, Global Change Biology, 2057–2064, doi: 10.1111/j.1365-2486.2005.01077.x.

Kandeler E., Mosier A.R., Morgan J.A., Milchunas D.G. , King J.Y. , Rudolph S., Tscherko D. (2006). Response of soil microbial biomass and enzyme activities to the transient elevation of carbon dioxide in a semi-arid grassland. Soil Biology & Biochemistry 38 (2006) 2448–2460.

Kant S., Bi Y. M., Rothstein S.J. (2010). Understanding plant response to nitrogen limitation for the improvement of crop nitrogen use efficiency, Journal of Experimental Botany, Pages 1-11 doi:10.1093/jxb/erq297.

Kelly J.J., Peterson E., Winkelman J., Walter T.J., Rier S.T., Tuchman N.C. (2013). Elevated atmospheric CO2 impacts abundance and diversity of nitrogen cycling functional genes in soil. Microbial Ecology 65(2):394-404.

Kessel C.V., Horwath W. R., Hartwig U., Harris D., LÜscher, A. (2000), Net soil carbon input under ambient and elevated CO_2 concentrations: isotopic evidence after 4 years. Global Change Biology, 6: 435–444. doi:10.1046/j.1365-2486.2000.00318.x.

Kimball, B. A., Lamorte, R. L., Peresta, G. J., Mauney, J. R., Lewin, K. F. and Hendrey, G. R. (1992). Weather, Soils, Cultural-Practices, and Cotton Growth Data from the 1989 Face Experiment in Ibsnat Format. Critical Reviews in Plant Sciences 2-3, 11: 271-308.

Long, S. P., Nie, G. Y., Baker, N. R., Drake, B. G., Farage, P. K., Hendrey, G. and Lewin, K. H. (1992). The Implications of Concurrent Increases in Temperature, CO_2 and Tropospheric O_3 for Terrestrial C_3 Photosynthesis. Photosynthesis Research 1, 34: 108-108.

Marzorati M., Wittebolle L., Boon N., Daffonchio D., Verstraete W. (2008) How to get more out of molecular fingerprints: practical tools for microbial ecology. Environ Microbiol. 10: 1571-1581.

Mauney, J. R., Lewin, K. F., Hendrey, G. R. and Kimball, B. A. (1992). Growth and Yield of Cotton Exposed to Free-Air CO_2 Enrichment (Face). Critical Reviews in Plant Sciences 2-3, 11: 213-222.

Minz D., Ofek M., Hadar Y. (2013). Plant rhizosphere microbial communities, The prokaryotes, 56-84doi : 10.1007/978-3-642-30123-0.

Morgan J.A.W., Bending G.D., White P.J. (2005) Biological costs and benefits to plant–microbe interactions in the rhizosphere. J Exp Bot 56:1729–1739.

Nannipieri P., Paul E. (2009).The chemical and functional characterization of soil N and its biotic components, Soil Biology & Biochemistry 41 (2009) 2357–2369.

Nelson D.M., Cann I.K.O., Mackie R.I. (2010). Response of Archaeal Communities in the Rhizosphere of Maize and Soybean to Elevated Atmospheric CO_2 Concentrations, PLoS ONE 5(12): e15897. doi:10.1371/journal.pone.0015897.

Pereira E.I.P., Chung H., Scow K., Sadowsky M.J., Kessel C.V., and Six J.(2012). Soil nitrogen transformations under elevated atmospheric CO2 and O3 during the soybean growing season, Environmental Pollution 159(2): 401–407.

Philippot L., Raaijmakers J.M. , Lemanceau P., Putten W.H.V. (2013). Going back to the roots: the microbial ecology of the rhizosphere, *Nature Reviews Microbiology* 11, 789–799.

Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., Peplies J., Glöckner F.O. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41 (D1);D590-D596.

Raup, D.M., Crick, R.E., 1979. Measurment of faunal similarity in paleontology. Journal of Paleontology 53, 1213–1227.

Rotthauwe J.H., Witzel K.P., Liesack W. (1997), The ammonia monooxygenase structural gene amoA as a functional marker: Molecular fine-scale analysis of natural Ammonia-Oxidizing populations, Applied and Environmental Microbiology,63(12) 4704-4712.

Sakurai M., Suzuki K., Onodera M., Shinano T., Osaki M. (2007). Analysis of bacterial communities in soil by PCR–DGGE targeting protease genes,Soil Biology and Biochemistry 39 (11); 2777-2784.

Schlesinger W. H. , Bernhardt E. S. , DeLucia E. H. , Ellsworth D. S. , Finzi A. C. , Hendrey G. R. , Hofmockel K. S. , Lichter J. , Matamala R. , Moore D. , Oren R. , Pippen J. S. , Thomas R. B. (2006). The Duke Forest FACE Experiment: CO_2Enrichment of a Loblolly Pine Forest. Ecological Studies, Chapter Managed Ecosystems and CO_2, Volume 187, pp 197-212.

Schloss, P.D., Westcott S.L., Ryabin T., Hall J.R., Hartmann M., Holliste E.B., Lesniewski R.A., Oakley B.B., Parks D.H., Robinson C.J., Sahl J.W., Stres B., Thallinger G.G., Van Horn D.J., Weber C.F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol, 2009. 75(23):7537-41.

Sims G.K., Wander M.M. (2002). Proteolytic activity under nitrogen or sulfur limitation, Applied Soil Ecology, Volume 19, Issue 3, Pages 217–221.

Xuexia Y., Xiangui L., Haiyan C., Rui Y., Huayong Z., Junli H., Jianguo Z. (2006). Effects of elevated atmospheric CO2 on soil enzyme activities at different nitrogen application treatments, Acta Ecologica Sinica, 2006, 26(1), 48−53.

Zhang X., Liu W., Schloter M., Zhang G., Chen Q., Huang J., Li L., Elser J.J., Han X. (2013). Response of the Abundance of Key Soil Microbial Nitrogen-Cycling Genes to Multi-Factorial Global Changes, PLoS ONE 8(10): e76500. doi:10.1371/journal.pone.0076500.

4. Microbial activities and protease expression in Barley rhizosphere during dark and light photoperiods: a metatranscriptome study

4.1 Abstract

Rhizosphere is the zone of nutrient exchange between plant and microbes and microbial activity in this zone gets influenced by the nutritional outputs from the plants, that depends on plant physiology and metabolism. Plants have different biochemical physiology during day, in exposure to light and in night during dark. The objective of this study was to investigate activities of the microbial populations during night and day in Barley rhizosphere. We studied active microbial pathways and protease expression in Barley rhizosphere (Hordeum vulgare) using metatranscriptome approach in a green -house experiment. Night samples showed significantly higher ($p<0.05$) activities for metabolism of nucleotides, vitamins and cofactors, Carbohydrate metabolism and amino acid metabolism pathways. Actinobacteria were among the most active organisms in both night and day samples. At class level higher significant activity was noticed for *Alphaproteobacteria* and *Betaproteobacteria* and at order level *Burkholderiales*, *Rhizobiales*, *Xanthomonadales* and *Sphilngomonadales* were more active in night samples than in day samples. Subtilisin type serine peptidases of family S8 and metallo-endopeptidases of family M4 peptidases, both had higher expression in night samples than in day samples. *Micromonosporacea* were dominant source of both types of studied proteases in all samples. This study report higher microbial activities in night samples than day samples for pathways related to C and N metabolism in Barley rhizosphere soil.

4.2 Introduction

In soil microbial diversity, Biomass, activities and functions are influenced by physiology of plants, their metabolism and distribution on a landscape (Garbeva et al., 2006, Helal and Sauerback, 2007, Lamb et al 2011, Lange et al 2015). Any factor affecting plant physiology has an impact on microbial population in soil and this effect on soil microbes is because of root exudates as confirmed by labeled isotope based studies (Bottner et al., 1999). Furthermore, effect due to plants is more pronounced in the rhizosphere and it exhibits higher microbial activities than bulk soil (Badalucco et al. 1996, Brzostek et al., 2013). Plants exhibit different metabolisms during day and night. During day plants carry out photosynthesis to fix atmospheric C and assimilate it in organic forms and during night plants utilize assimilated C. This difference in day and night metabolism leads to different natures of root exudates and higher exudation rates were reported during day than in night in Barley roots (*Hordeum vulgaris*) (Liljeroth et al. 1990). Because of different natures of root exudates fluctuations in soil redox potential has been observed resulting in diurnal shifts in microbial activities in rhizosphere soil (Nikolausz et al. 2008).

There are different approaches used to study the microbial population in soil. Most popular approaches are based on 16S gene based identification of bacteria (Janssen 2006, Vasilieadis et al. 2012). Identification of microbial populations based on 16S gene is not enough as it only gives information about communities present in soil but no information about functions. Another approach involves study of functional genes involved in important processes controlling N and C metabolism (Wallenstein and Vilgalys 2005, Hai et al. 2009, Wang et al. 2012, Yergeau et al. 2007). In another advanced improvement to study soil microbial populations metagenome is studied that include study of all the genes from a

microbial community in an environmental sample (Daniel 2005, Fierer et al. 2012). Study of DNA based approaches have a limitation that it only gives a picture about the potential of fuctions and not real activity. For assessing real activities, RNA based approaches are often more reliable. A metatranscriptome that involves study of messenger RNA (mRNA) gives an expression profile of a soil microbial population from an environmental sample.

Present study was carried out to determine the rhizospheric microbial population activities and protease expression in response to the changes in plant metabolism during dark and light photoperiods. To investigate the rhizospheric microbial activities during dark and light photoperiods we have used a metatranscriptome approach on rhizosphere soil from Barley plant (*Hordeum vulgare*, Barke cultivar) in a green-house experiment. Samples were collected 1 hour before and 1 hour after sunrise. Obtained metatranscriptome sequences were BLASTed against different databases to study differences in active pathways in night and day samples and to study the active microbial populations. Two types of proteases; Subtilisin type serine endopeptidases (S8 peptidase family) and Metallo-endopeptidases of M4 peptidase family were studied.

4.3 Materials and methods

Experimental setup: Soil for this experiment was collected from an agricultural farm located in Schyern in Germany (48°N, 11°E) and had organic C content of 16±0.8 mg g^{-1}, N content of 1.76±0.1 mg g^{-1}, a C/N-ratio of 9.0±0.4 and a pH value ($CaCl_2$) of 6.6±0.1. Soil was sieved through 2 mm mesh and was adjusted to 50% water holding capacity and were filled in plastic pots of dimensions 9x9x11 cm were filled with 700 kg soil.

After an equilibration time of 7 days, Barley seedlings were sowed in each of the pots.Prior to sowing Barley in pots, healthy Barley seeds were selected, were surface sterilized with hypochlorite solution and after thorough rinsing with sterile distilled water, seeds were germinated in sterile petrilplates at 37 ° C for 2-3 days. 8 healthy looking sprouts were selected and were planted one in each of the prepared pots. Pots were then kept in green house under alternate light and dark periods. Dark periods were maintained from 20.00 to 6.00 hrs., at temperature of 18 °C, whereas light period were maintained from 6.00 to 20.00 hrs., at temperature of 20 °C. Pots were regularly watered 3 days a week with 100 ml of distilled water.

Sampling, nucleic acid extraction and metatranscriptome library preparation: Samples were collected on 20th day that corresponds to tillering stage of Barley growth. Samples were collected in triplicates. Night samples were named N1, N2 and N3, and day samples were named D1, D2 and D3. Night sampling was carried out at 4.00 am 1 hour before sunrise and day time sampling was carried out at 6.00 am 1 hour after sunrise. Samples were collected by uprooting the plants from pots carefully so as to prevent breaking of roots. Roots were shaken well to remove loosely adhered soil, roots along with soil that was left after shaking was immediately preserved in Lifeguard solution (MO BIO) and stored in -20°C till extraction.

Prior to nucleic extraction, RNA lifeguard solution was completely removed by centrifuging at 1500 g for 5 minutes. 0.5 g soil was weighed in BIO101 lysing matrix tubes (MP Biomedicals) and nucleic acids were extracted as per Griffith's protocol (Griffith et al., 2000), with slight modification where 10µl/ml ß-Mercaptoethanol was added to phosphate buffer before use.DNA-RNA co-extract pellet thus obtained was dissolved in 50 µl of deionized-DEPC treated water. DNA was depleted using MOBIO DNase Max kit. Purity of RNA was

confirmed by 16S gene universal PCR using primers 27f/1492r (Lane 1991). The reaction mixture contained 2.5 µl of each Top Taq Buffer (Qiagen), Coral solution (Qiagen), 1.5 µl of Q-solution (Qiagen), 2nM dNTPs, 0.2 µM of each primer and 1 µl of template in a 25 µl reaction. Quality and integrity was checked on Agilent Bioanalyzer 2100 using Agilent RNA 6000 Nano Kit (Agilent technoilogies). RNA was quantified using Quant-iT™ RiboGreen® RNA Assay Kit (ThermoFischer Scientific) and fluorescence was measured on a Gemini EM microplate reader (Molecular devices, USA) using SoftMax Pro data acquisition and analysis software V5.0 (Molecular devices, USA)

rRNA was depleted using Ribo-Zero kit (Bacteria)-Low input(Epibio). Depleted mRNA was fragmented, reverse transcripted and metatranscriptome libraries were prepared using Script-seq complete Kit (Bacteria)-Low input (Epibio) as per manufacturer's instructions except where only 0.5 µl of primers were used for library preparation instead of 1 µl. Quality of libraries prepared were checked on Agilent Bioanalyzer 2100 using Agilent High Sensitivity DNA kit as per manufacturer's instructions.

Sequencing and sequence analysis: Metatranscriptome libraries were sequenced on a MiSeq Illumina sequencing system (Illumina). Obtained sequences were in FASTQ form. Reads were first concatenated, then adapters were removed, followed by removal of contaminants by deconseq (Schmeider and Edwards, 2011). Reads were then converted to FASTA format and ribosomal RNA was removed using SortMeRNA (Kopylova et al. 2012). Non-ribosomal RNA was then BLASTed against NCBI (Sayers et al. 2008) and KEGG (Kanehisa and Goto, 2000) databases to get an overview of the taxonomy and expressed functions. Later analyses were performed on MEGAN (Huson et al. 2007). For other statistical analyses R was used. For protease analysis sequences were BLASTed against MEROPS database.

4.4 Results

Most active pathways

Sequences were annotated at different levels of pathways against KEGG database. Figure 1a shows an overview of most active pathways in day and night samples at level 3. Metabolisms of nucleotides, vitamins and cofactors, carbohydrates were significantly higher ($p<0.05$) in night samples than in day samples. At level 4 among most active pathways significant differences were observed for pyruvate, Pyrimidine and Purine metabolism, Citrate cycle and prokaryotic Carbon fixation pathway (Figure 1b). Most pathways with high activity were related to Carbon and N metabolism, genetic information processing and nucleotide metabolism.

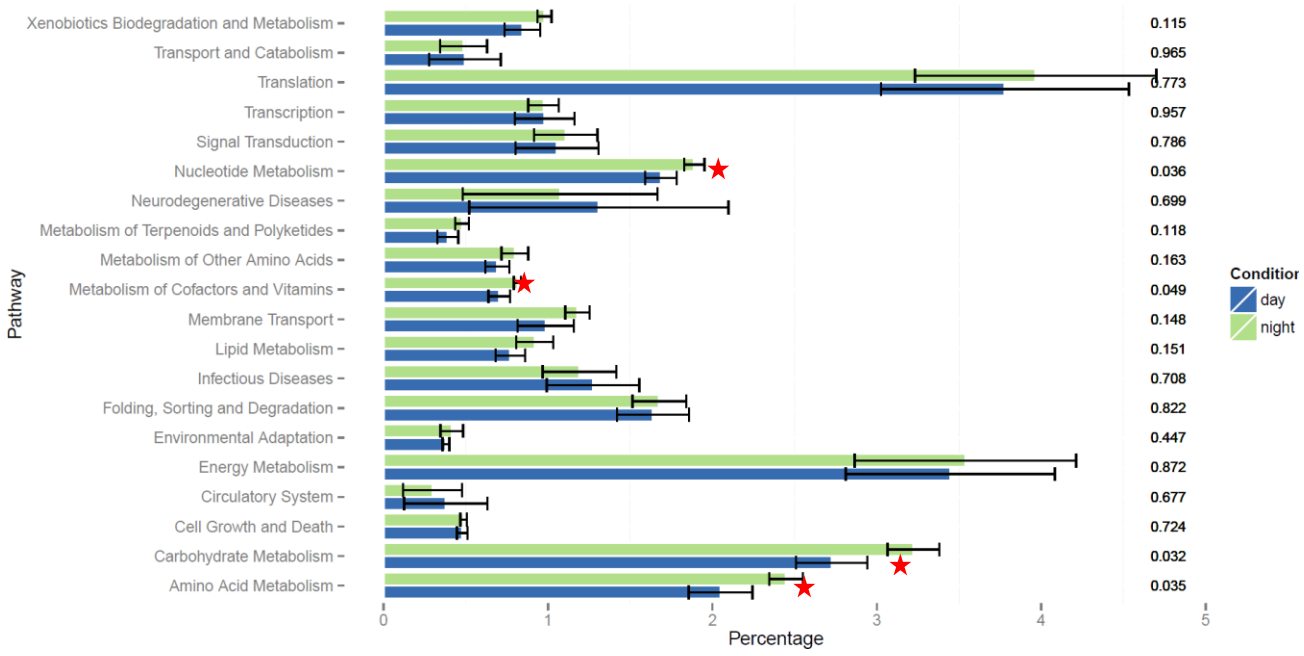


Figure 1a. Most active pathways at level 3 of KEGG database

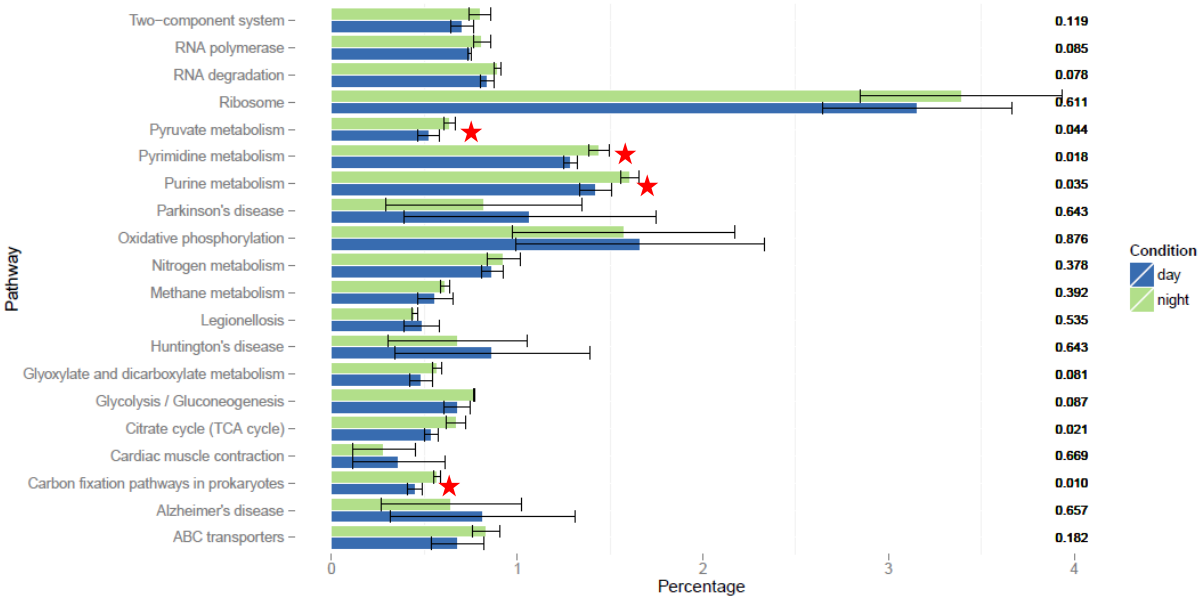


Figure 1b. Most active pathways at level 4 of KEGG database

Activity of microbes

At taxonomic level, highest activity was observed for *Actinobacteria* which play an important role in organic matter decomposition, but these organisms didn't show any significant differences in day and night samples. Apart from actinobacteria, most active organisms were those who plays active role in C and N cycling especially those involved in N fixation like alph*apr*oteobacteria and bet*apr*oteobacteria (Figure 1a), among *alphaproteobacteria* it was *Rhizobiales*, *Burkholderiales* among *betaproteobacteria* (Figure 2b)which were most abundant, both of these are well known for their roles in N-fixation. These results indicate a higher N fixation activity in night samples. At family, genus, kingdom, phylum, species and superkingdom level differences between day and night samples were not significant.

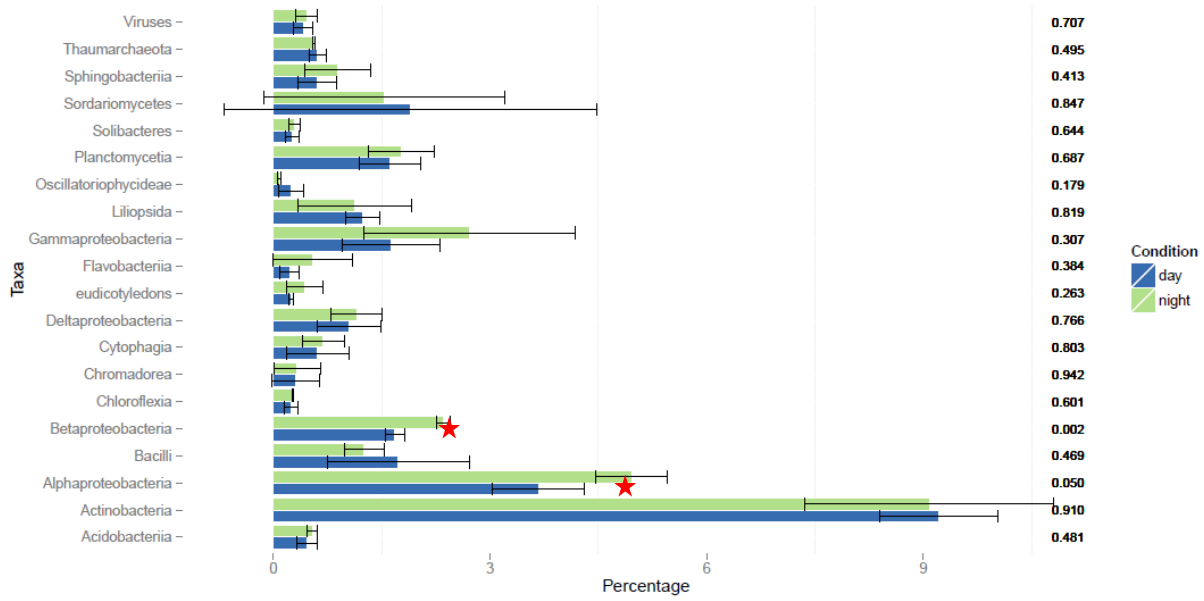


Figure 2a. Most active organisms at taxonomic class level

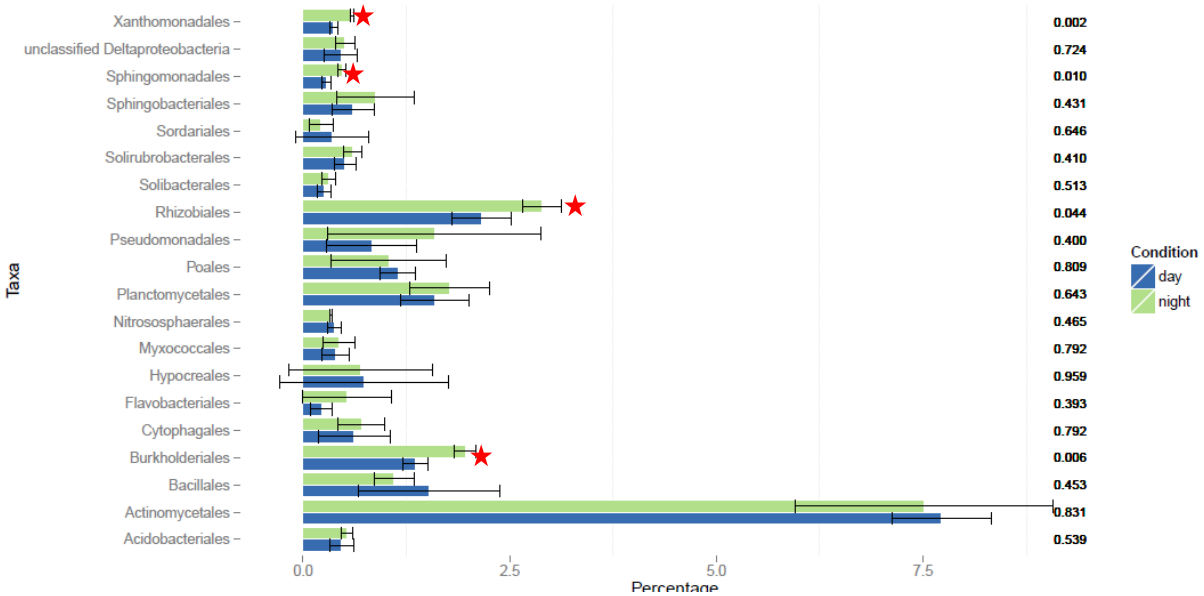


Figure 2b. Most active organisms at order level

Expression of proteases

S8 peptidase family consist of subtilisin type serine endopeptidases and M4 family consist of extracellular metalloendopeptidase. We observed a significant higher S8 and M4 peptidase expression in night samples than in day samples (Figure 3). We looked at most abundant 23

orders and families for S8 peptidases and 13 most abundant orders and 21 most abundant families for M4 peptidases. At order level most important contributors were Actinomycetales in both S8 and M4 peptidase families, contributing 31.80 % of all identified S8 sequences and 49.49% of all identified M4 sequences(Figure 4a and 4b). Micromonosporaceae that belongs to Actinomycetales order are most important contributor at taxonomic family level, contributing 9.86% of total sequences from S8 family and 13.35 % of total sequences from M4 family (Figure 5a and 5b).

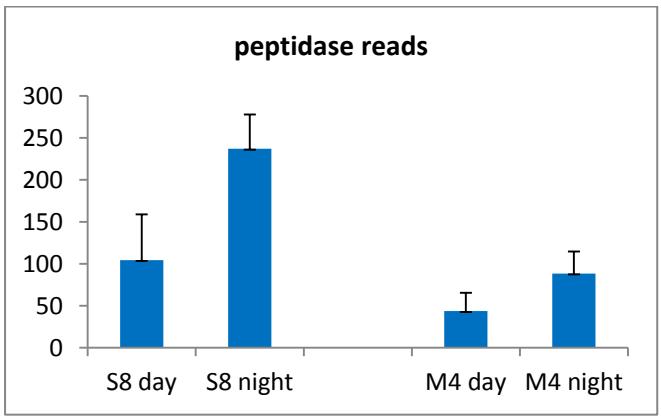


Figure 3 Expression of S8 and M4 peptidases in day and night samples

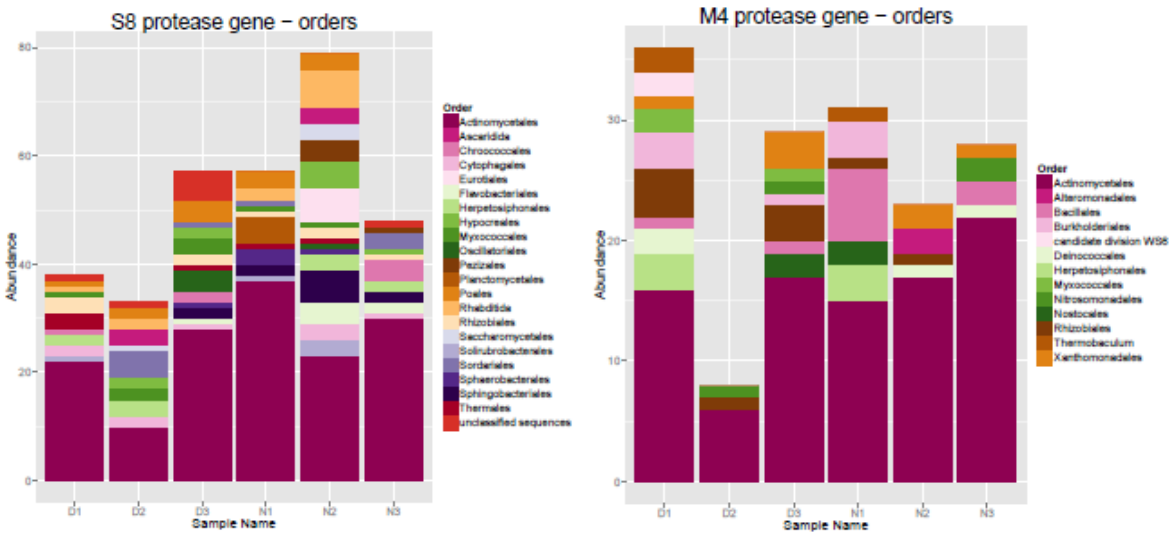


Figure 4 Abundance of best hits and order level for (a) S8 protease and (b) M4 protease

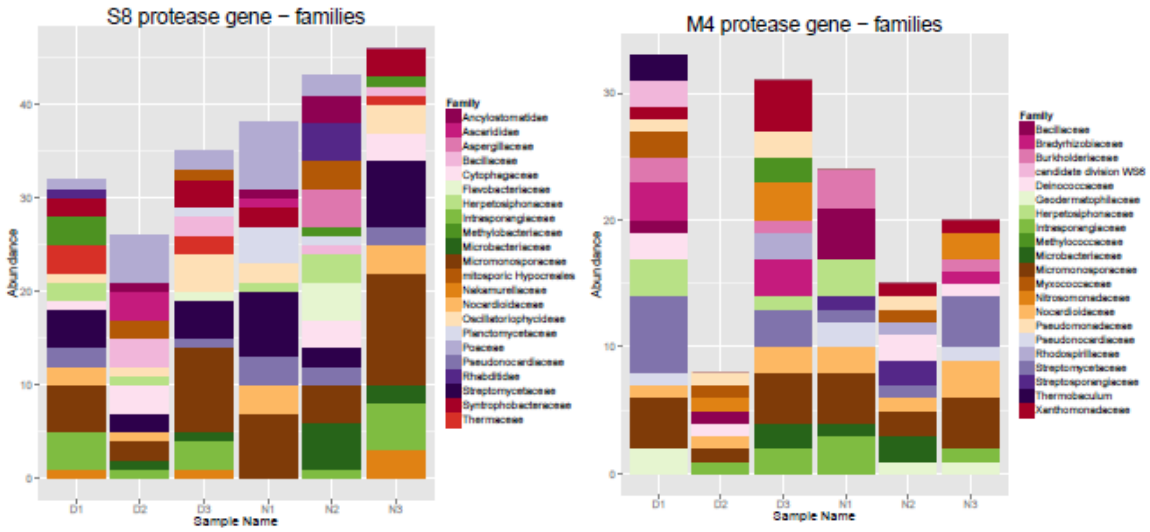


Figure 5 Abundance of best hits family level for (a) S8 protease and (b) M4 protease

4.5 Discussion

We noticed a higher activity related to Carbohydrate metabolism in night samples (Figure 1a) and particularly prokaryotic carbon fixation pathways (Figure 1b). Higher relative humidity during night favors higher activity of microbiota and higher respiration (Medina and Zelwer, 1972). Harris and Van Bavel, 1957 reported highest plant respiration at 4.00 am in cotton, tobacco and corn plants. During day there is a negative priming effect due to photosynthesis resulting in reduced soil organic matter (SOM) mineralization (Kuzyakov and Cheng, 2004), in night due to respiration . Another factor that controls soil priming effect is temperature and higher temperature leads to a positive priming effect (Li et al. 2011), but since in our experiment the difference between day and night temperature was only 2°C, we assume that the effect is due to photosynthesis. Higher respiration results in higher metabolism of stored carbon sources in cells and higher priming effect. Apart from carbon metabolism we noticed higher activities in pathways involved in metabolism of amino acids, co-factors, vitamins and nucleotides in night samples. Metabolism of amino acids, co-factors, vitamins and nucleotides indicate adaption to specific environmental adaption

strategy (Gianoulis et al. 2009). This indicates that during dark hours rhizospheric microbes change environmental strategies to adapt to changed conditions to nutrient input in soil and root exudates from plants.

Actinobacteria were among the most abundant class of bacteria in both night and day samples (Figure2a). Though we didn't notice significant differences in abundance between day and night samples previous studies by Bulgarelli et al. 2015 also confirmed high abundance of *Actinobacteria* and *Proteobacteria* in Barley Rhizosphere, and *Actinobacteria* are known for their multiple roles in soil (Aislabie et al. 2013). Night samples showed higher expression for orders *Rhizobiales*, *Xanthomonadale*s, *Sphingomonodales* and *Burkholderiales*. Members of order *Rhizobiales* (class *Alphaproteobacteria*) have an important role in Biological N fixation (Carvalho et al. 2010, Jones 2015). *Xanthomonadales* order (class *Gammaproteobacteria*) consists of well known phyto pathogens that use different mechanisms to infect plants (Alfano and Colmer, 1996, Sanchez 2011). Bacteria from order *Sphingomonadales* can utilize C from both root exudates and soil organic matter (Lakshmanan et al. 2014). *Burkholderials* are known as N fixers (Caballero-Mellado et al. 2007)in addition to this they also have role in suppressing soil borne diseases induced by mixed hay-cropping system (Benitez and Gardener 2009). Higher activity of these organisms in night samples should be a result of more favorable conditions during dark for the type of activities these microbes are involved in. During day plants produce more Reactive oxygen species (ROS) due to higher temperature and photosynthesis (Michelet and Krieger-Liszkay, 2012). In night time mitochondria are main sources of ROS production due to respiration (Rhoads et al. 2006), unlike during day photorespiration and photosynthesis are main source of ROS production. Nitrogenase is an oxygen sensitive enzyme and this could be a possible

explanation for the higher activities of microbial taxons involved in N fixation in night samples.

Bacteria secrete M4 metallo-peptidases extensively in soil and sediments and have been proposed to play an important role in environmental protein degradation (Rao et al. 1998, Wu and Chen 2011). Similarly S8 peptidases are produced by a wide range of organisms and are present in various soil environments (Tripathi and Sowdhamini, 2008). We have observed a lower expression in day samples compared to night samples for both these protease families. Previous studies in grass seedlings have reported that during day oxidative denaturation of proteins takes place due to action of ROS from plants takes place and in night oxidized proteins from plant roots and rhizospheric microbes carry out proteolysis, and amino acids thus produced are used by plants for nutrition and growth (White et al. 2015). Furthermore higher expression of proteases in night samples could be explained again by the bacterial mining for SOM decomposition due to priming effect.

In summary microbes in Barley rhizosphere responds to different inputs from plants by changing their metabolic strategies. During night they exhibited higher activity of pathways that involve carbohydrate metabolism, prokaryotic C fixation, and metabolisms of amino acids, vitamins, co-factors and nucleotides. Activities of certain organisms particularly those involved in N-fixation which use oxygen sensitive nitrogenase complex, gets suppressed during day time when there is higher production of ROS resulting from photosynthesis, photorespiration and cellular respiration, compared to higher activity during night when only cellular respiration is the source of ROS production. Higher expression of protease enzymes were also reported in night samples. This study will help us in understanding the

diurnal response of rhizospheric microbes to photosynthetic and non-photosynthetic phases of plants in day and night.

References

Aislabie J., Deslippe J., Dymond J. (2013). Soil microbes and their contribution to soil services, in Ecosystem services in New Zealand: conditions and trends, 143- 161.

Alfano J.R., Collmer A. (1996). Bacterial pathogens in plants: Life up against the wall, Plant Cell 8:1683-1698.

Badalucco L., Kuikman P. J. , Nannipleri P. (1999). Protease and deaminase activities in wheat rhizosphere and their relation to bacterial and protozoan populations, Biology and Fertility of Soils, 23(2) pp 99-104.

Benítez Mar M.S., Gardener B.B. M. (2009). Linking Sequence to Function in Soil Bacteria: Sequence-Directed Isolation of Novel Bacteria Contributing to Soil borne Plant Disease Suppression, Applied Environmental Microbiology ,75 ; 4915-924.

Bottner P., Pansu M., Sallih Z. (1999). Plant and Soil, Volume 216,Issue1, pp 15-25.

Brzostek E.R., Greco A., Drake J.E., Finzi A.C. (2013). Root carbon inputs to the rhizosphere stimulate extracellular enzyme activity and increase nitrogen availability in temperate forest soils, Biogeochemistry, 115(1) pp 65-76.

Bulgarelli D., Garrido-Oter R., Münch P.C., Weiman A., Dröge J., Pan Y., McHardy A.C., Schulze-Lefert P. (2015). Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley, Cell Host and Microbe 17 (3); 392-403.

Carvalho F.M., Souza R.C., Barcellos F.G., Hungria M., Vasconcelos A.T.R. (2010). Genomic and evolutionary comparisons of diazotrophic and pathogenic bacteria of the order Rhizobiales, *BMC Microbiology*, 10:37 doi:10.1186/1471-2180-10-37.

Daniel R. (2005). The metagenomics of soil, *Nature Reviews Microbiology* 3, 470-478.

Fierer N., Leff J.W., Adams B.J., Nielsen U.N., Bates S.T.**,** Lauber C.L., Owens S., Gilbert J.A.**,** Wall D.H., Caporaso J.G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes, Proceedings of National Academy of Sciences of the United States if America 109(52): 21390–21395.

Garbeva P., Postma J., van Veen J.A., van Elsas J.D. (2006). Effect of above-ground plant species on soil microbial community structure and its impact on suppression of Rhizoctonia solani AG3, Environmental Microbiology, 8(2):233-46.

Gianoulis T. A., Raes J., Patel P. V., Bjornson R., Korbel J. O., Letunic I., Yamada T., Paccanaro A., Jensen L.J., Snyder M., Bork P., Gerstein M. B. (2009). Quantifying environmental adaptation of metabolic pathways in metagenomics. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(5), 1374–1379. http://doi.org/10.1073/pnas.0808022106.

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(12):5488-91.

Gschwendtner S., Reichmann M., Müller M., Radl V. , Munch Jean C. , Schloter M. (2010). Abundance of bacterial genes encoding for proteases and chitinases in the rhizosphere of three different potato cultivars, Biology and Fertility of soils 46 (6), pp 649-652.

Hai B., Ndeye H.D., Sall S., Haesler F., Schauss K., Bonzi M., Assigbetse K., Chotte J.L., Munch J.C., Schloter M. (2009). Quantification of Key Genes Steering the Microbial Nitrogen Cycle in the Rhizosphere of Sorghum Cultivars in Tropical Agroecosystems, Applied and Environmental Microbiology 75 (15); 4993-5000.

Harris D.G., Van Bavel C.H.M. (1957). Root respiration of tobacco, corn and cotton plants, Agronomy Journal, 49, 182-184.

Helal, H. M. and Sauerbeck, D. (1986), Effect of plant roots on carbon metabolism of soil microbial biomass. Z. Pflanzenernaehr. Bodenk., 149: 181–188.
doi: 10.1002/jpln.19861490205

Huson D.H., Auch A.F., Qi J., and Schuster S.C. (2007). MEGAN analysis of metagenomic data, Genome Research 17(3): 377–386. doi: 10.1101/gr.5969107.

Janssen P.H. (2006).Identifying the Dominant Soil Bacterial Taxa in Libraries of 16S rRNA and 16S rRNA Genes, Applied Environmental Microbiology 72(3): 1719–1728.

Jones, R. T. (2015) A Comprehensive Survey of Soil Rhizobiales Diversity Using High-Throughput DNA Sequencing, in Biological Nitrogen Fixation (ed F. J. de Bruijn), John Wiley & Sons, Inc, Hoboken, NJ, USA. doi: 10.1002/9781119053095.ch76.

Kanehisa M., Goto S. (2000). KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Research 28(1):27-30.

Kopylova E., Noé L. and Touzet H., (2012). “SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data", *Bioinformatics* doi: 10.1093/bioinformatics/bts611.

Kuzyakov Y., Cheng W. (2004). Photosynthesis controls of CO_2 efflux from maize rhizosphere, Plant and Soil, 263 (1);85-99.

Lakshmanan V., Selvaraj G., Bais H.P. (2014). Functional Soil Microbiome: Belowground Solutions to an Aboveground Problem, Plant Physiology 166 (2); 689-700.

Lamb E.G., Kennedy N., Siciliano S.D. (2011). Effects of plant species richness and evenness on soil microbial community diversity and function, Plant and Soil (2011) 338:483–495.

Lane, D. J. (1991). "16S/23S rRNA sequencing," in *Nucleic Acid Techniques in Bacterial Systematics*, eds E. Stackebrandt and M. Goodfellow (Chichester: John Wiley & Sons), 115–175.

Lange M., Eisenhauer N., Sierra C.A. , Bessler H., Engels C., Griffiths R.I. , Mellado-Vázquez P.G. , Malik A.A. , Roy J., Scheu S., Steinbeiss S., Thomson B.C. , Trumbore S.E., Gleixner G. (2015). Plant diversity increases soil microbial activity and soil carbon storage, *Nature Communications* 6, Article number:6707 doi:10.1038/ncomms7707.

Li Z., Wang X., Zhang R., Zhang J., Tian C. (2011).Contrasting diurnal variations in soil organic carbon decomposition and root respiration due to a hysteresis effect with soil temperature in a *Gossypium* s. (cotton) plantation, Plant and soil 343(1); 347-355.

Liljeroth E., Bååth E., Mathiasson I., Lundborg T. (1990). Root exudation and rhizoplane bacterial abundance of barley (*Hordeum vulgare*L.) in relation to nitrogen fertilization and root growth, Plant and Soil,127(1); pp 81-89.

Medina E., Zelwer, M. (1972), in Papers from a Symposium on Tropical Ecology with an Emphasis on Organic Productivity (ed. Golley, P. M.), Univ. Georgia, Athens, pp. 245–269.

Michelet, L., Krieger-Liszkay, A. (2012). Reactive oxygen intermediates produced by photosynthetic electron transport are enhanced in short-day grown plants. Biochimica et biophysica Acta (BBA)-Bioenergetics 1817: 1306–1313.

Nikolausz M., Kappelmeyer U., Székely A., Rusznyák A., Márialigeti K., Kastner M. (2008). Diurnal redox fluctuation and microbial activity in the rhizosphere of wetland plants. European Journal of Soil Biology, 44, 324-333.

Rao M.B., Tanksale A.M., Ghatge M.S., Deshpande V.V. (1998) Molecular and Biotechnological Aspects of Microbial Proteases. Microbiol Mol Biol Rev 62: 597–629.

Rhoads D. M., Umbach A. L., Subbaiah C. C., Siedow J. N. (2006). Mitochondrial Reactive Oxygen Species. Contribution to Oxidative Stress and Interorganellar Signaling. *Plant Physiology*, *141*(2), 357–366. http://doi.org/10.1104/pp.106.079129.

Sanchez C. (2011).Microbial ecology: Bacteria reinforce plant defences, *Nature Reviews Microbiology* 9, 483 doi:10.1038/nrmicro2598.

Sayers E.W., Barrett T., Benson D.A., Bryant S.H., Canese K., Chetvernin V.,Church D.M., DiCuccio M., Edgar R., Federhen S., Feolo M., Geer L.Y., Helmberg W., Kapustin Y., Landsman D., Lipman D.J., Madden T.L., Maglott D.R.,Miller V., Mizrachi I., Ostell J., Pruitt K.D., Schuler G.D., Sequeira E.,Sherry S.T., Shumway M., Sirotkin K., Souvorov A., Starchenko G., Tatusova T.A., Wagner L., Yaschenko E., Ye J. (2009). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2009 Jan;37(Database issue):D5-15. Epub 2008 Oct 21.

Schmieder R., Edwards R., (2012). Fast identification and removal of sequence contamination from genomic and metagenomic datasets, PLoS One 6(3):e17288. doi: 10.1371/journal.pone.0017288.

Tripathi L, Sowdhamini R (2008) Genome-wide survey of prokaryotic serine proteases: analysis of distribution and domain architectures of five serine protease families in prokaryotes. BMC Genomics 9:549.

Vasileiadis S., Puglisi E., Arena M., Cappa F., Cocconcelli P.S., Trevisan M.(2012). Soil bacterial diversity screening using single 16S rRNA gene V regions coupled with multi-million read generating sequencing technologies,PLoS One 7(8): doi: 10.1371/journal.pone.0042671

Wallenstein M.D., Vilgalys R.J. (2005). Quantitative analyses of nitrogen cycling genes in soils, Pedobiologia 49 (6); 665-672.

Wang H., He Z., Lu Z., Zhou J., Nostrand J.D.V., Xu X., Zhang Z. (2012). Genetic Linkage of Soil Carbon Pools and Microbial Functions in Subtropical Freshwater Wetlands in Response to Experimental Warming, Applied and Environmental Microbiology 78 (21); 7652-7661.

White J.F., Chen Q., Torres M.S., Mattera R., Irizarry I., Tadych M., Bergen M. (2015). Collaboration between grass seedlings and rhizobacteria to scavenge organic nitrogen in soils, AoB Plants 7 : plu093doi: 10.1093/aobpla/plu093

Wu J.W., Chen X.L. (2011). Extracellular metalloproteases from bacteria. Appl Microbiol Biotechnol 92: 253–262.

Yergeau E., Kang S., He Z., Zhou J., Kowalchuk G.A.(2007).Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect, ISME Journal 1(2):163-79.

5. Conclusions and perspectives

Studies in this thesis contribute to our knowledge of soil proteases and the response of proteolytic genes triggered by various biotic and environmental factors. In all the three approaches, we have observed a strong influence of root exudates on microbial functional diversity and abundance. Inherent plant NUE that is governed mainly by a plant's genetic makeup has potential to trigger an increased protease enzyme activity, diversity and abundance of genes encoding proteases, apart from proteases we noticed a positive trigger for other enzyme activities (Chitinase and Urease activities) involved in soil N-minerlaization (as observed in chapter 2). Through rhizobox approach we noticed that the significant influence on proteolytic genes was limited to rhizosphere soil and bulk soil wasn't affected much, neither in terms of abundance nor in diversity. So we conclude that the root exudates strongly affect microbial proteolytic genes and microbial proteolytic communities in soil, and a higher NUE plant often trigger-on proteolytic genes for a higher extracellular protease enzyme activity in soil. A large population of bacteria contributoring towards protease gene pool in soil are also plant growth promoting rhizo-bacterias (PGPRs). This confirms the integral role of proteases in maintaining soil health and fertility of soil.

On one hand we have used a biotic factor as a variable, in another approach we used an environmental factor that is CO_2 as a variable to look for its effect on abundance in rhizosphere and bulk soil. On raising CO_2 to 550 ppm concentration in air using a FACE setup we noticed a shift in abundance of proteolytic genes and also of amoA genes. We looked for diversity and abundance of *apr* protease gene and bacterial amoA gene along with study of microbial diversity by Illumina sequencing of 16S genes. Shifts in mentioned two genes and microbial diversity was visible not only in rhizosphere soil but also in bulk soil and so we

assume that here eCO_2 has a direct effect on soil, apart from plant mediated effect. Effect of plants is not limited due to inherent plant metabolism but plant physiology also affects proteases, particularly in rhizosphere soil. In chapter 3 we observed that peptidase families S8 and M4 are higher expressed in night than in day in Barley rhizosphere soil in a green-house experiment. Overall results indicate that in soil, activity, diversity and expression of protease encoding genes are affected by a plethora of different factors and plants have a strong influence on controlling activity, diversity and abundance of these genes. This research has unravelled the so far unknown response, diversity and distribution of protease encoding genes in soil and this information can be used to improve soil health and fertility.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisors Prof. Paolo Nannipieri and Prof. Giacomo Pietramellara for their continuous support of my Ph.D study and related research, for their patience, motivation, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisors and mentors for my Ph.D study. Their encouragement kept me motivated and enthusiastic for my research.

Besides my advisor, I would like to thank Prof. Giancarlo Renella , Dr. Judith Ascher, Dr. Maria Teresa Ceccherini and Dr Laura Giagnoni from DISPAA department, for their insightful comments and encouragement and for helping me widen my research from various perspectives. Without their precious support it would not be possible to conduct this research.

I thank our technician Dr. Paola Arfaioli for all the support and fun we have had in the last three years. Also I thank mybestest friend from the department Dr. Federica Borgogni who was always a very supportive person and a source of inspiration and I learnt how to stand strong from her.

It is my genuine pleasure to thanks Prof. Michael Schloter and Dr. Anne Schoeler from Helmholtz Zentrum in Munich, for the collaboration and providing a wonderful opportunity to work on a very innvoative project. I would like to also thanks Dr. Edoardo Puglisi for helping me a lot in Bioinformatics data analysis.

My heartful thanks to Marie Curie ITN action project 'TRAINBIODIVERSE'for financing my PhD and providing me an excellent platform in my research career.

Last but not the least, I would like to thank my family: my parents, my brother and my all close and dear ones for supporting me throughout writing this thesis and my life in general. Special thanks to both my grandparents without you people it would have been impossible to finish this thesis.