FEMS Microbiology Ecology Advance Access published June 24, 2015

Metagenomic analyses reveal no differences in genes involved in cellulose degradation under different tillage treatments

Maria de Vries[1,4], Anne Schöler[1,4,5], Julia Ertl[1], Zhuofei Xu[2,3], Michael Schloter[1]

3rd revised version (June19th, 2015)

[1] Environmental Genomics, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg

[2] University of Copenhagen,Microbiology, Universitetsparken 15, 2100 København Ø

[3] Present address:KU Leuven, Stem Cell Biology and Embryology, O&N IV Herestraat 49 - box 804,3000 Leuven

[4] equal contributions

[5] corresponding author

Phone: +49 (0)89 3187 3048, Fax: +49 (0)89 3187 3376, anne.schoeler@helmholtz-muenchen.de

Abstract

Incorporation of plant litter is a frequent agricultural practice to increase nutrient availability in soil and heavily relies on the activity of cellulose degrading microorganisms. Here we address the question how different tillage treatments affect soil microbial communities and their cellulose degrading potential in a long-term agricultural experiment. To identify potential differences in microbial taxonomy and functionality, we generated six soil metagenomes of conventional (CT) and reduced (RT) tillage-treated topsoil samples, which differed in their potential extracellular cellulolytic activity as well as microbial biomass.

Taxonomic analysis of metagenomic data revealed few differences between RT and CT and a dominance of *Proteobacteria* and *Actinobacteria*, whereas eukaryotic phyla were not prevalent. Prediction of cellulolytic enzymes revealed glycoside hydrolase families 1, 3, 5, 94, auxiliary activity family 8 and carbohydrate binding module 2 as the most abundant in soil. These were annotated mainly to the phyla of *Proteobacteria*, *Actinobacteria* and *Bacteroidetes*.

These results suggest that the observed higher cellulolytic activity in RT soils can be explained by a higher microbial biomass or changed expression levels but not by shifts in the soil microbiome. Overall this study reveals stability of soil microbial communities and cellulolytic gene composition under the investigated tillage treatments.

Introduction

The most abundant organic polymer on earth is cellulose. As a key component of plant cell walls it is highly abundant in all plants mostly in combination with hemicellulose and xylan (Varner & Louis 1989). The degradation of cellulose is an essential ecosystem service and it is of key value to gain more quantitative and qualitative understanding of the role of cellulose breakdown within the carbon cycle and in a changing environment. Cellulases are in general hydrolytic and divided into three major types: β-1,4-endoglucanases (EC3.2.1.4), β-1,4-exoglucanases and β-glucosidases (EC3.2.1.21) (Lynd et al. 2002). Exoglucanases are further divided into cellobiohydrolases or, cellobiosidases (EC3.2.1.91, EC3.2.1.176) and cellodextrinases (EC3.2.1.74). These cellulases differ in their binding and cutting sites on the cellulose fibers or oligomers. Besides hydrolysis, other catalytic modes of action to cut cellulose chains have been identified, namely oxidoreduction, mediated by the enzyme cellobiose dehydrogenase (CBH, EC 1.1.99.18, (Langston et al. 2011)) or induced by a quinone- or glycopeptide-mediated Fenton reaction (Baldrian & Valášková 2008), and phosphorolytic degradation (EC 2.4.1.20, EC 2.4.1.49, (Reichenbecher et al. 1997)). Cellulases generally consist of a catalytic module, classified into glycoside hydrolase- (GH-) families, and often harbour a carbohydrate binding module (CBM), whereas cellulases with oxidoreductive catalytic modules are classified as auxiliary activities- (AA-) family of proteins. The classification into modules is based on amino acid sequence similarity (Henrissat 1991) and a comprehensive description of these module families can be found in the CAZy database (Carbohydrate-active enzymes database; www.cazy.org, (Lombard et al. 2014)).

Different agricultural practices like tillage and fertilization strategies influence world-wide carbon cycles in soil (Lal 2004; Schimel 1995; Schlesinger & Andrews 2000). Tillage, a key component of modern agriculture, disperses added nutrients throughout the soil and aerates the soil, but also affects soil aggregates

(Abdollahi et al. 2014). There is a strong ongoing debate on how different tillage treatments impact not only yields, but also soil nutrients and diversity and activity of soil microbial communities. To clarify this question several long term agricultural experiments have been established to assess the adaptation of soils to different tillage treatments. In this context it was shown that reduced tillage (RT) leads to a higher carbon content and higher microbial biomass in the topsoil compared to conventional tillage (CT) (van Groenigen et al. 2011; Küstermann et al. 2013; Alvarez et al. 1995; Kandeler et al. 1999).

However, it is currently unclear how different tillage treatments influence the phylogenetic structure of the soil microbial community, in particular those which drive the degradation of polymers, like cellulose, which are major constituents of plant residues used to improve soil quality. Only a few studies have compared conventional tillage to no-tillage systems and these studies have detected several differences in taxonomy and functionality of the associated microbial communities (Souza et al. 2013; Carbonetto et al. 2014). For example in a long-term (26 years) experiment, deep tillage practice led to a lower microbial diversity in the top 10 cm of the soil (Silva et al. 2013). However, no in-depth analysis focussing on microbes catalysing carbohydrate degradation has been performed so far in agricultural systems.

Thus, in this study we assess the impact of different tillage treatments on microbial communities which drive cellulose degradation in soil. Samples were taken from the topsoil (0-10 cm) of a long-term field experiment, where a split plot-based study has been established with conventional (CT; 25 cm working depth) and reduced tillage (RT; 8 cm working depth) treatments (Meyer-Aurich et al. 2009; Küstermann et al. 2013). One month before sampling, corn had been harvested and the remaining plant litter had been incorporated into the soil with the different tillage treatments. Due to the different tillage depths the amount of plant residues and consequently the amount of straw in the topsoil differed between the two treatments (Tebrügge et al. 1999). To assess the influence of different tillage treatments on the enzymatic activity of key cellulolytic enzymes in the soil, the potential enzymatic activities of β-glucosidase and cellobiohydolase were measured using methylumbeliferone-labeled substrates.

We explored the influence of long-term conventional and reduced tillage on the taxonomic and functional diversity of soil microbial communities, which are involved in cellulose degradation. As it is well accepted that cellulases occur ubiquitously in all kingdoms (http://www.cazy.org, Lombard et al. 2014) and show a high genetic diversity, we employed a shotgun sequencing approach to identify the different cellulases present and the microbes harbouring these genes in a qualitative (diversity) and semi-quantitative manner. The analysed top soils differed in their amount of total organic carbon (Küstermann et al. 2013). Therefore, we hypothesized to find a higher relative abundance and diversity of genes coding for cellulases in soil samples from RT compared to CT.

Experimental Procedures

Site description and soil sampling

Soil was sampled from an agricultural field experiment, established in 1992 at a research farm in Scheyern, 40km north of Munich (Germany) (Meyer-Aurich et al. 2009; Küstermann et al. 2013). From this field experiment, two tillage treatments were analysed: conventional tillage with moldboard plough (25 cm working depth) and reduced tillage with rotary harrow (8 cm working depth). The experiment has been set up with three independent replicates (split plot design). Ten-year-average precipitation and temperature are respectively 792.3 mm and 8.3 °C (2000-2010). The soil is a Luvisol with a pH of 6.3, consisting of 2.2% coarse sand, 17.0% fine sand, 55.4% silt and 25.4% clay.

The soil was sampled in November 2012, one month after corn harvest and incorporation of the plant residues into the soil. Of every plot a composite sample of five soil cores to a depth of 10 cm was taken with a soil auger of 5 cm diameter, and sieved with a 3-mm sieve. Afterwards, a subset of the soil was mixed and immediately stored on dry ice and subsequently at -80 °C before DNA extraction. The remaining soil was stored at 4 °C.

Enzymatic activity assays, biomass and chemical measurements

Potential enzymatic activity was measured one day after sampling of soil stored at 4 °C. Potential cellulose degradation activity of the microbial community was assessed by adding methylumbeliferone-complexed (MU-)β-D-glucoside and MU-β-cellobioside to soil solutions (Pritsch et al. 2005). The optimum substrate concentration (C_{opt}) and incubation time (T_{opt}) for degradation of both substrates were determined during a pilot experiment; C_{opt}=800µM MU-ß-cellobioside and 500µM MU-ß-D-glucoside, T_{opt}= 120 min for both. Fluorescence measurement after incubation of the substrate with soil solution was done with a spectrophotometer at excitation wavelength λ=365 nm and emission wavelength of λ=450 nm. Fluorescence values were corrected for soil-autofluorescence or fluorescence inhibition by soil particles (quenching) and background noise.

For measurement of microbial carbon, water extractable organic carbon (WEOC) was determined before and after chloroform fumigation (Joergensen 1996). For determination of WEOC, soils were mixed with 0.01M $CaCl_2$ in a 1:5 ratio, incubated in an overhead shaker for 40 min at room temperature, filtered, and stored at -20 °C until measurement with the Dimatoc 100 (DimatecAnalysentechnik GmbH, Germany).

For measurement of total carbon (TC) and nitrogen (TN) soil was dried at 40 °C for ca. 5 days and ground to powder shaking in a tissue lyzer at 30 Hz for 3-9 min. 20-25 mg of ground soil powder was weighed in duplicate, wrapped in aluminium and subjected to elemental analysis using an Elementar Vario EL III instrument in combustion mode.

DNA isolation & Pyrosequencing

Soil DNA was extracted from 300 mg of frozen soil samples using the DNA-isolation kit ‘Genomic DNA from soil’ NucleoSpin Soil Kit (Macherey– Nagel, Germany) according to the protocol of the manufacturer. DNA was stored at -20 °C until further processing. Library preparation was performed according to the Roche protocol “Rapid Library Preparation Method Manual” using Roche Molecular Identifier (MID) adapters as barcodes. Amplification of DNA via emulsion PCR was done according to the Roche protocol “emPCR Method Manual - Lib-L Large Volume (LV)”. Pyrosequencing was carried out on the Genome Sequencing (GS) FLX+ instrument, using a GS FLX Titanium sequencing kit XL+. Image- and signal processing was performed by the provided Roche software. The sequences are deposited in SRA under the BioProject ID: PRJNA235154.

Real-time PCR assays

Quantitative real-time PCR (qPCR) was performed using primers for amplification of the bacterial 16S ribosomal gene (Bach et al. 2002) and the fungal Internal Transcribed Spacer (ITS) regions (White et al. 1990). The reaction mixtures (total volume 25 µl) for quantification of the 16S ribosomal gene consisted of: 12.5 µl Power SYBR Green master mix (Life Technologies), 5 pmol of each primer, 0.5 µl 3% bovine serum albumin (BSA) and 2 µl soil DNA template. The reaction mixtures (total volume 25 µl) for quantification of the ITS gene consisted of: 12.5 µl Power SYBR Green master mix (Life Technologies), 10 pmol of each primer and 2 µl soil DNA template. For quantification, standard curves were calculated using serial dilutions (10^2 to 10^6 gene copies μl^{-1}) of a plasmid containing *Fusarium oxysporum* DNA (for fungal qPCR) or *Pseudomonas putida* (for bacterial qPCR). PCR detection limit was assessed to 10 gene copies μl^{-1}. In order to prevent PCR inhibition, the optimal dilution for each amplification assay was determined by dilution series of DNA extracts (data not shown). The qPCR assays were performed in 96-well plates. All PCR runs started with a hot start at 95˚C for 10 minutes, then either 40 cycles of 94˚C for 30 s, 55˚C for 30 s and 72˚C for 30 s for ITS samples or 40 cycles of 95°C for 45 s, 58°C for 45 s and 72°C for 45 s for 16S qPCR. Finally melting curve analyses were carried out for all samples by a final cycle of 95˚C for 15 s, 60˚C for 30 s and 95˚C for 15 s. The amplification efficiency was calculated from the formula $\mathrm{Eff} = [10^{(-1/\mathrm{slope})} - 1]*100$, and resulted in the following efficiencies: 16S: 102% and ITS: 87%.

Analysis of sequencing data and prediction of cellulolytic enzymes

Sff-files were separated by MID adapters, using 454/Roche SFF Tools. Trimming of reads was carried out using a modified DynamicTrim (Cox et al. 2010) as supplied by MG-RAST (Meyer et al. 2008) using the following settings: minimum Phred score= 15, maximum number of bases below minimum Phred score=5, minimum read length=50. Remnant adapter sequences and duplicated reads were removed using Biopieces (www.biopieces.org) and cd-hit (Fu et al. 2012), respectively. For taxonomic and functional annotation, filtered reads were blasted against both the National Center for Biotechnology Information (NCBI) non-redundant protein database and the Kyoto Encyclopedia of Genes and Genomes (KEGG)-database (June 2011) using Diamond (Buchfink et al. 2014) as well as against the Silva SSU-database using Blastn (expect (e)-value threshold= 10^{-4}). Mapping of the top 25 blast results (i.e. hits with lowest e-value) to taxonomic and functional annotations was carried out with the MetaGenome ANalyzer (MEGAN, Version 5.2.3) software (http://ab.inf.uni-tuebingen.de/software/megan5/). During the MEGAN analysis the following settings were used: Min support=1, Min score=50, Top%=10, Min-Complexity Filter=0. Visualization of data was performed in R (R Core Team 2013). To obtain a relative abundance of reads, the number of reads was divided by the total number of filtered reads per sample and multiplied by 100.

For prediction of catalytic modules (GH and AA) and carbohydrate binding modules (CBM) in the reads, protein open-reading frames were predicted using FragGeneScan (Rho et al. 2010) and translated to protein sequences. Different families of catalytic modules and CBMs harboured by the key enzymes in cellulose degradation β-1,4-endoglucanases, β-1,4-exoglucanases, β-glucosidases, cellodextrin phosphorylases and cellobiose dehydrogenases, were selected from the CAZy database (www.cazy.org (Lombard et al. 2014)). Protein Hidden Markov Model (HMM)-logos can be used to scan protein sequences using hmmscan (Eddy 2011). For the selected families, several HMM-logos are available in the Protein family (Pfam) A 26.0 database (Punta et al. 2012) and in the DataBase for automated Carbohydrate-active enzyme Annotation (dbCAN, Yin Y et al., 2012), but can also be personally built using hmmbuild (which is contained within the HMMER version 3.0 (March 2010), packaged together with hmmscan; http://hmmer.org/) and an alignment-file with sequences containing the corresponding module domain as input (alignment files were generated using sequences from CAZy for each corresponding family).

Because the selected module families also contain a varying amount of enzymes other than cellulases, we decided to validate the specificity of these HMM-logos for cellulases by scanning them against a list of positive or negative cellulase sequences. These lists have been assembled by downloading CAZy entries for all glycoside hydrolase, carbohydrate binding and auxiliary activity families that contain cellulolytic proteins. A list of 1283 entries was then classified as positive, having a matching EC number (as mentioned in the introduction) and a list of 1390 entries as negative where the EC number did not match. Of the positive sequences, 662 were eukaryotic. For the families AA8 and AA9, entries that mentioned activity on cellulose in the name were included in the positive list. Taking all Genbank Ids for these positive and negative entries, the corresponding protein sequences were downloaded from NCBI. Finally, these lists were scanned with the three sets of HMM-logos (from Pfam A, dbCAN or personally built) and the HMM-logos that annotated most of the positive sequences (see supplementary **Table S1** in bold) were used to scan the translated protein sequences from the metagenomes (e-value threshold= 10^{-4}).

To limit the number of false positive predictions we refined the search for cellulases by performing an additional quality control step, i.e. blasting all HMM-annotated protein sequences against the above-mentioned positive sequences database using blastp (e-value threshold= 10^{-5}). To obtain a relative abundance of reads, the number of reads was divided by the total number of filtered reads per sample and multiplied by 100,000.

For taxonomic annotation, all reads with a predicted cellulolytic module, i.e. after scanning against the HMM-logos and the positive sequences database, were blasted against the NCBI non-redundant protein database using Diamond (minscore=50) and mapped with MEGAN5 (parameters as previously announced).

Statistical analysis of sequencing data and diversity analysis

To detect global differences between the metagenomes, principal component analysis was performed on the relative order count data of metagenomic reads annotated to the NCBI non-redundant protein database and the KEGG database and analysed with MEGAN5. Significant differences of soil measurements and annotated read counts between CT and RT soils were determined by unpaired t-test statistics, and adjusted using the Bonferroni correction for metagenomic data, using R (R Core Team 2013). Before testing for differences between soil parameter-measurements, QQ-plots and Shapiro-Wilk tests for normal distribution were performed. For the PCA-plot, the data was square-root transformed, as the QQ-plot showed no normal distribution. Differences were regarded significant when the adjusted *p*-value was smaller than 5 % (P<0.05). In addition, only taxa or pathways with a relative abundance of at least 0.05% in one of the replicates were included. Due to the low amount of predicted cellulases, no abundance cut-off was used to detect possible differences between tillage treatments. Shannon indices were calculated using the vegan package in R (Oksanen et al. 2013).

Results

Microbial biomass and enzymatic activities

As expected, we found higher total soil carbon and nitrogen contents under conditions of reduced tillage (C=1.62 ± 0.13%, N= 0.17 ± 0.01%) compared to conventional tillage treatment (C=1.08 ± 0.07%, N= 0.12 ± 0.01%, *P*=0.004 respectively *P*=0.004). In addition, reduced tillage leads to significantly higher microbial biomass compared to soil under conventional tillage, as measured by carbon analysis (0.28 ± 0.03 respectively 0.16 ± 0.02 mg g^{-1} dry weight (DW), *P*=0.004, **Fig. 1a**) and by qPCR of both bacterial (8.37 ± 0.49 x 10^{10} respectively 6.82 ± 0.19 x 10^{10} copies g^{-1} DW, *P*=0.007) and fungal DNA (5.58 ± 1.03 x 10^{8} respectively 3.40 ± 0.40 x 10^{8} copies g^{-1} DW, *P*=0.027). The high ratio of bacterial 16S rRNA genes to fungal ITS genes was not influenced by the treatments. Potential extracellular enzyme activity measurements showed that CT results in a lower activity than RT for both β-glucosidase (278.7 ± 40.9 respectively 545.1 ± 86.1nmol MU h^{-1} g^{-1} DW,*P*=0.049) and cellobiohydrolase (5.0 ± 1.6 respectively 28.8 ± 2.7nmolMU h^{-1} g^{-1} DW, *P*=0.002) (**Fig. 1b and 1c**).

Taxonomic and functional annotation of metagenomes

Shotgun sequencing of the six DNA libraries (3 independent replicates per tillage treatment) resulted in an overall amount of 0.5 Gigabases of data, which corresponds to a mean of 157,106 filtered reads per replicate with an average length of 410 bps (**Table S2**).To account for the different number of reads between the biological replicates, relative abundances per replicate were calculated. Taxonomic annotation of the metagenomes was performed by blasting all filtered reads against the Silva SSU database and assigning the taxonomic rank with MEGAN5. Due to the low amount of ribosomal sequences in the metagenomes (0.12%), only a classification at kingdom level was performed and revealed a dominance of Bacteria (86.15%) followed by Eukaryota (13.39%) and Archaea (0.45%) (data not shown). 61% of the reads could be taxonomically annotated to at least super-kingdom level with the NCBI non-redundant protein database. Therefore, all further taxonomic annotations refer to the NCBI database. At super-kingdom level the majority of reads was mapped to Bacteria (98.03%) followed by Eukaryota (1.08%), Archaea (0.73%) and Viruses (0.05%). Exploratory rarefaction analysis at each taxonomic level revealed an acceptable coverage of diversity on the order level for all six samples (**Fig. 2).** Taxonomic analysis showed that all soil metagenomes are dominated by *Proteobacteria* and *Actinobacteria* that together make up for 26.3% of all reads (**Fig. S1a**). Further prevalent bacterial phyla were *Bacteroidetes* (3.3%), *Acidobacteria* (3.3%), *Verrucomicrobia* (1.7%), *Gemmatimonadetes* (1.4%), *Planctomycetes* (1.3%) and *Chloroflexi* (1.1%). At the order level *Actinomycetales*, *Rhizobiales*, *Myxococcales*, *Burkholderiales* and *Planctomycetales* accounted for 13.5% of all annotated reads (**Fig. S1b**).

The principal component analysis performed on the relative order count data from the metagenome reads (**Fig. S2**) revealed no clear clustering of replicates from conventional and reduced tilled soils. In accordance, no significant change in alpha-diversity (Shannon indices: conventional tillage: 3.58±0.02, reduced tillage: 3.55±0.04) between the metagenomes was detected. Significantly different abundances of taxonomically annotated metagenome reads between CT and RT soils are shown in Table S3. Genes annotated to originate from the order *Ktedonobacteraceae* were more abundant in conventional tillage-treated soil compared to reduced tillage-treated soil (*P*=0.003).

For functional annotation, the metagenomes were mapped to the KEGG database, to which 35.21 % of the reads could be annotated, and visualized with MEGAN5. Two-component system, Purine metabolism, ABC transporters, Nitrogen metabolism and Pyrimidine metabolism were among the most abundant pathways and account for 6.14 % of all reads (**Fig. S3**). Soils under conventional tillage contained significantly more genes involved in Carbohydrate metabolism (*P*=0.008), Xenobiotics Biodegradation and metabolism (*P*=0.0006), Arachidonic acid metabolism (*P*=0.006), Tyrosine Metabolism (*P*=0.010) and Drug Metabolism (*P*=0.0006, **Table S3**).

Prediction of cellulolytic enzymes from metagenomes of conventional and reduced tillage-treated soils

When focusing on the enzymes involved in cellulose degradation within the starch and sucrose metabolism pathway of KEGG, the β-glucosidase (EC 3.2.1.21) and endoglucanase (EC 3.2.1.4) were most abundant, with 270 reads (mean relative abundance 0.071% ± 0.008) for conventional and 410 reads (mean relative abundance 0.073% ± 0.003) for reduced tillage annotated to β-glucosidase. 123 reads (mean relative abundance 0.032% ± 0.004) for conventional and 186 reads (mean relative abundance 0.033% ± 0.001) for reduced tillage were annotated to endoglucanase. These numbers revealed no significant difference between conventional and reduced tillage. Much fewer reads were annotated to exoglucanase (EC 3.2.1.91, CT: 0.001%±0.0006 and RT: 0.002%±0.0018) and cellobiose phosphorylase (EC 2.4.1.20, CT: 0.002%±0.0014 and RT: 0.001%±0.0005) and none to cellobiose dehydrogenase (EC 1.1.99.18). In the KEGG database there is no reference KEGG Orthology groups for cellobiosidase (EC 3.2.1.176) or cellodextrinase (EC 3.2.1.74).

More extensive screening of the metagenomes for cellulases was performed using hmmscan. To identify the most sensitive HMMs for cellulolytic proteins we benchmarked publically available as well as personally built HMMs against a set of positive and negative CAZy proteins (see Experimental Procedures). This revealed that most HMMs detected also many negative CAZy proteins. Using the most sensitive HMMs (**Table S1** in bold**),** 5,906 reads could be annotated to a catalytic or carbohydrate binding module. To reduce the number of false positive predictions using hmmscan, an additional filtering step was performed by scanning the predicted cellulases against the same set of positive CAZy proteins (**Table S4**). After this quality control step 2,021 reads remained with a predicted catalytic or carbohydrate binding function, which corresponds to 0.214 % of the total amount of metagenome reads. The percentage of total annotated cellulases was not different between treatments. A large variety of catalytic modules and CBMs that are involved in cellulose degradation were found in the metagenomes of this agricultural soil; 18 glycoside hydrolases, 4 auxiliary activities and 14 carbohydrate binding modules (**Fig. 3**). After quality control, most hits were found for GH1, GH3, GH94, AA8 and CBM2. Statistical analysis revealed significantly more reads annotated to CBM11 in conventional tillage-treated soil than in reduced tillage-treated soil (*P*=0.020).

Conversely, reduced tillage-treated soil contained more reads annotated to GH48 than conventionally tilled soil (*P*=0.032).

The taxonomic affiliation of the pooled cellulase-annotated sequences (from Figure 3) is shown in **Fig. 4** on phylum and order level**,** and reflects the overall abundance of phyla in the metagenomes (**Fig. S1a**). 18.56% of all cellulase reads mapped to the *Proteobacteria,* 11.23% to *Actinobacteria* and fewer to *Bacteroidetes* (8.61%), *Cyanobacteria* (2.23%), *Gemmatimonadetes* (2.18%), *Verrucomicrobia* (2.13%), *Acidobacteria* (1.73%), *Firmicutes* (1.53%) and *Chloroflexi* (1.53%). *Actinobacteria, Bacteroidetes, Cyanobacteria* and *Firmicutes* show a high abundance here, whereas they are not so abundant in the total metagenome (respectively 7.1%, 3.3%, 0.7% and 0.8%of total reads, **Fig. S1a**). Results on order level support the importance of the role which *Actinobacteria* (*Actinomycetales*), *Proteobacteria* (*Rhizobiales* (α), *Burkholderiales* (β), *Xanthomonadales* (γ), *Myxococcales* (δ)), *Bacteroidetes* (*Cytophagales, Sphingobacteriales*) and *Verrucomicrobia* (*Verrucomicrobiales*) play in cellulose degradation.Of the reads that were predicted to contain cellulolytic modules 0.71% mapped to fungi (*Ascomycota* and *Basidiomycota*).

Finally, the taxonomic affiliation for the individual most abundant cellulolytic families (GH1, 3, 5, 94, AA8 and CBM2) and the difference between conventional and reduced tillage is shown separately in Figure S4. The low number of reads leads to a high variation in the amount of annotated reads among replicates and treatments.

Discussion

Relative abundance and diversity of genes coding for cellulolytic enzymes

We predicted a multitude of cellulolytic enzymes in our study, which reflects the complex nature of cellulose degradation in soil. The functional annotation of cellulase genes agrees with other studies searching for catalytic modules and CBMs in aerobic environments, where also genes of GH family 3, 5 and 9 have been identified (DeAngelis et al. 2010; Anderson et al. 2012; Duan & Feng 2010; Nyyssönen et al. 2013). This is also true for their taxonomic annotation, as cellulolytic microorganisms have been found in the phyla of *Proteobacteria, Firmicutes, Bacteroidetes* and *Actinobacteria* (Woo et al. 2013; Anderson et al. 2012;

Nyyssönen et al. 2013; Berlemont & Martiny 2013). Also the order of *Actinomycetales* contains well known cellulose degraders (Větrovský et al. 2014; Solans & Vobis 2003), as does the order of *Myxococcales* (Reichenbach et al. 2006).

To our knowledge, a similar refining analysis on HMM-annotated reads as presented here (**Fig. 3**) has not previously been done. We find a highly variable specificity of each HMM-motif for cellulolytic function, which ranges from 0.0 to 100.0 % of specificity (**Table S4**). Our results show that such an analysis is necessary to get a more exact prediction of gene function, as results can depend heavily on the nature of the HMM. However, sequence similarity scoring may not be the optimal method to annotate diverging protein-coding sequences. To define the actual activity of the predicted cellulases, other approaches like cloning and subsequent functional characterization, transcriptomics and proteomics are necessary in addition.

To further assess the question what percentage of predicted cellulases can be more reliably annotated, we compared the sequences predicted to harbour cellulolytic domains (GH and AA) to the sequences that mapped to KEGG modules. This revealed that 19.5% of all sequences predicted by the HMMs before filtering and 28.8% after filtering with the curated database were annotated by the KEGG database as well. This indicates that at least a part of the identified sequences is potentially functional in the soil.

In the overall metagenome, 1.08 % (NCBI nucleotide database) or 13.39 % (Silva SSU database) was annotated to eukaryotes. Indeed, the qPCR on the soil DNA showed a very low abundance of fungal ITS sequences, especially compared to bacterial 16S sequences. These results are comparable to earlier measured ITS-sequence abundance in agricultural soil (Ng et al. 2012). In accordance with the overall metagenome, 0.93% of the predicted cellulases in the metagenome were annotated to eukaryotes (0.71% was annotated to fungi) when using the NCBI non-redundant protein database. However, 12% of the predicted cellulases were annotated to a CAZy module which contain mostly eukaryotic sequences (GH7, CBM1 and 49, AA3, 8 and 9). Discrepancies between the annotation results of different databases (NCBI versus Silva or NCBI versus CAZy) clearly show the database biases towards sequences from often-studied organisms. In a high-throughput approach, numbers of reads annotated to taxa or functions are only realistic if a sufficient database-representation of the corresponding genes exists. Unfortunately, this is not the case, as databases are biased towards culturable bacteria (Wooley et al. 2010) and care should be taken when assessing results based on similarity to database contents. In addition, the discrepancy between the annotated abundance of fungi by the Silva SSU database and the qPCR results presented here, might be explained by the detection of a different fungal sequence (i.e. 18S in the Silva SSU database, and ITS by the PCR assay). Besides database annotation biases, fungal genes contain many intronic sequences, which require long reads for accurate annotation. A transcriptomic approach might resolve the hidden contribution of fungi in the soil metabolic function, as is evidenced by other studies targeting soil transcripts (Kellner et al. 2010; Baldrian et al. 2012).

The low amount of fungi annotated in the overall metagenome and the cellulase-annotated sequences is striking when one considers the role of fungi in biomass degradation. In forest soils, fungal cellulases are known to be important for cellulose degradation (Bailey et al. 2002; Buée et al. 2009). It is however known that agricultural soils harbour lower amounts of fungi than undisturbed soils, as tillage disrupts soil fungal communities (Miller & Lodge 1997). Furthermore, the yearly application of mineral fertilizers and fungicides negatively influences the amount of fungi in the sampled soil.

Influence of tillage treatment on soil metagenomes

We hypothesized to find a higher abundance and diversity of cellulase genes in RT soil compared to CT soil, based on the assumption of higher substrate availability under RT. Few significant differences were found between the metagenomes, e.g. genes related to GH48 were exclusively found in RT soil, whereas CT soil harboured more CBM11 genes (**Fig. 3**). GH48-family proteins are generally harboured by bacteria and anaerobic fungi, and are always present in cellulosomes. Furthermore, GH48 cellulases are considered the key component of various cellulolytic systems (Devillard et al. 2004; Olson et al. 2010; Izquierdo et al. 2010). CBM11-family proteins are only known to be present in bacterial cellulases. However, so far not much is known of the function of cellulase diversity or specific groups of cellulases in the degradation process in soil. Future research will have to elucidate the specific contribution of each family. Another finding is that more reads were annotated to the class *Ktedonobacteria* in CT soil than in RT soil (**Table S3**). *Ktedonobacteria* are gram-positive, aerobic organisms that produce branched vegetative mycelia (Yokota et al. 2012). However, not much is known of their physiology and growth conditions or ecological importance. Using the KEGG database for functional annotation of the metagenomes we detect significantly more reads for carbohydrate metabolism and biodegradation of xenobiotics in CT soil compared to RT soil. This is an interesting finding but difficult to interpret as there are many downstream pathways and none of them differ between tillage treatments.

Most detected differences were however based on low read counts, questioning their ecological relevance. For this a metagenomic assessment with higher resolution or more biological replicates can be helpful to perform more powerful statistics. In general, this in-depth analysis of soil metagenomes from RT and CT soils reveals no global differences in taxonomic and functional content, and none in alpha-diversity.

This is to our knowledge the first metagenomic analysis of the same agricultural soils under different tillage practices. It is therefore plausible that a change in tillage practice has far less impact on a functioning microbial community than a comparison of tillage versus no-tillage, for which several differences in microbiome structure and taxonomic composition were detected (Carbonetto et al. 2014; Souza et al. 2013). In this respect we can hypothesize that conditions below 10 cm in the RT soil in this study might resemble conditions under no-tillage. Indeed, Angers *et al*. showed that the surface layer and deeper soil layers can be differentially affected by tillage treatment (Angers et al. 1997). In the subsoil of RT, no extra carbon is incorporated. Accordingly, no differences in soil organic carbon between CT and RT were observed at deeper soil levels (18-25 cm) at this experimental site (Küstermann et al. 2013). To confirm whether the subsoil of RT is affected in a similar way as no-tillage soil, future research is required.

Our results indicate that tillage intensity does not strongly influence microbial community structure in the timeframe of 20 years, and implies that these microbial communities in agricultural soils are stable to some extent in response to long-term differences in tillage treatments. The impact of different tillage treatments might prove less selective for soil microbial community members than other environmental factors such as the regional climate. This theory is in agreement with a recent meta-analysis of reduced tillage systems in organic farming (Cooper, personal communication).

In contrast to the similarity between metagenomes, microbial biomass, fungal and bacterial signature gene copies and cellulase enzyme activity, which were determined per gram soil, were clearly lower in CT than RT topsoil. In agreement with these observations, a higher fungal biomass in reduced compared to conventionally tilled soil has been previously reported (Murugan et al. 2013; van Groenigen et al. 2010). Due to a difference in microbial biomass, a higher bacterial 16S- and fungal ITS-abundance and enzymatic activity in RT compared to CT soils might be an indirect result of tillage. Alternatively, the higher enzymatic activity might be due to a higher abundance of cellulase enzymes in RT soil.

Several other explanations for the discrepancy between the enzyme activity measurements and the metagenomic analysis exist. One potential explanation is that the soil microbiome might primarily react on differences in tillage treatments on a transcriptional level. The metagenomic approach does not reveal the active members of a given microbial community and it is likely that these will be different under conventional and reduced tillage practice. A metatranscriptomic or metaproteomic approach would be required to reveal the effect of tillage on the expression or translation of specific cellulase genes. The advantage of these approaches is that the active organisms are addressed directly, but the disadvantage is that such measurements can fluctuate considerably over time. Furthermore, it is known that sequencing depth influences the detection of rare or abundant species (Wooley et al. 2010), so it is conceivable that in this study many rare microbial species have been missed that contribute to cellulose degradation. However, if these species have a very low abundance, it is unlikely that their role in degradation is ecologically significant.

Another explanation for the observed discrepancy is the sampling time. The incorporation of plant litter leads to a high input of carbon into the soil. It is conceivable that at the time of sampling after one month, the whole microbial community might be profiting from an increased labile carbon pool and there would be little competition between microorganisms. This is in agreement with findings in a field study by Fu *et al.*, which showed that soil respiration was still increasing after 30 days of crop residue incorporation, both under no-tillage and conventional tillage (Fu et al. 2000). Sampling at a later time point (e.g. half a year) might reveal stronger differences between degradation potential of microbial communities of CT and RT soils due to the lower amount of carbon left in the topsoil of CT compared to RT.

Outlook

Cellulose degradation replenishes the labile carbon pool in soil and is an important ecosystem service. In this study we addressed the question how different tillage treatments affect cellulose degrading communities in soil in an agricultural setting. As expected, we identified a high diversity of cellulases in soil and several glycoside hydrolase families with high abundances, but did not find evidence that tillage treatment strongly influences cellulose degrading communities at the level of DNA.

We detect only few differences in the microbial communities of conventionally versus reduced tillage-treated soils as well as in the abundance of cellulolytic genes. The differences observed in potential enzymatic activities could not be explained by changes in diversity and might be driven by the increased microbial biomass present at RT sites or different expression patterns. However, as the present study is purely based on functional predictions, a confirmation of the results by more targeted methods that confirm the cellulolytic activities are needed, e.g. by using isotopic labelled cellulose as substrate and subsequent stable isotope probing analysis, or a transcriptomic approach further combined with degenerate primer-gene capture. Finally, further quantification of genomic or transcribed cellulases in response to different biotic and abiotic factors, as well as studies that improve our understanding on the dynamics of these microbial communities in time and space, will be a crucial step to better understand to what extent the large diversity of cellulolytic enzymes is relevant for efficient cellulose degradation.

Acknowledgments We thank Susanne Kublik and Marion Engel for their help to perform the sequencing run, Gisle A. Vestergaard for bioinformatics help, Gudrun Hufnagel and Franz Buegger for measuring carbon and nitrogen fractions, and Gerhard Welzl for advice on the statistical analysis.

Maria de Vries and Anne Schöler are supported by fellowships from the Marie Curie ITN TRAINBIODIVERSE network.

References

Abdollahi L, Munkholm LJ & Garbout a (2014) Tillage System and Cover Crop Effects on Soil Quality: II. Pore Characteristics. *Soil Sci. Soc. Am. J.* 78: 271–279.

Alvarez R, Diaz RA, Barberob N, Santanatogliaa OJ & Blotta L (1995) Soil organic carbon , microbial biomass and C02-C production from three tillage systems. *Soil Tillage Res.* 33: 17–28.

Anderson I, Abt B, Lykidis A, Klenk H-P, Kyrpides N & Ivanova N (2012) Genomics of aerobic cellulose utilization systems in actinobacteria. *PLoS One* 7: e39331.

Angers D a. et al. (1997) Impact of tillage practices on organic carbon and nitrogen storage in cool, humid soils of eastern Canada. *Soil Tillage Res.* 41: 191–201.

Bach HJ, Tomanova J, Schloter M & Munch JC (2002) Enumeration of total bacteria and bacteria with genes for proteolytic activity in pure cultures and in environmental samples by quantitative PCR mediated amplification. *J. Microbiol. Methods* 49: 235–245.

Bailey VL, Smith JL & Bolton H (2002) Fungal-to-bacterial ratios in soils investigated for enhanced C sequestration. *Soil Biol. Biochem.* 34: 997–1007.

Baldrian P et al. (2012) Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *ISME J.* 6: 248–258.

Baldrian P & Valášková V (2008) Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiol. Rev.* 32: 501–521.

Berlemont R & Martiny AC (2013) Phylogenetic distribution of potential cellulases in bacteria. *Appl. Environ. Microbiol.* 79: 1545–1554.

Buchfink B, Xie C & Huson DH (2014) Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12.

Buée M, Reich M, Murat C, Morin E, Nilsson RH, Uroz S & Martin F (2009) 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytol.* 184: 449–456.

Carbonetto B, Rascovan N, Álvarez R, Mentaberry A & Vázquez MP (2014) Structure, composition and metagenomic profile of soil microbiomes associated to agricultural land use and tillage systems in Argentine Pampas. *PLoS One* 9.

Cox MP, Peterson DA & Biggs PJ (2010) SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* 11: 485.

DeAngelis KM et al. (2010) Strategies for Enhancing the Effectiveness of Metagenomic-based Enzyme Discovery in Lignocellulolytic Microbial Communities. *BioEnergy Res.* 3: 146–158.

Devillard E, Goodheart DB, Karnati SKR, Bayer EA, Lamed R, Miron J, Nelson KE & Morrison M (2004) Ruminococcus albus 8 Mutants Defective in Cellulose Degradation Are Deficient in Two Processive

Endocellulases, Cel48A and Cel9B, Both of Which Possess a Novel Modular Architecture. *J. Bacteriol.* 186: 136–145.

Duan C-J & Feng J-X (2010) Mining metagenomes for novel cellulase genes. *Biotechnol. Lett.* 32: 1765–1775.

Eddy SR (2011) Accelerated Profile HMM Searches. *PLoS Comput. Biol.* 7: e1002195.

Fu L, Niu B, Zhu Z, Wu S & Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28: 3150–3152.

Fu S, Coleman DC, Schartz R, Potter R, Hendrix PF & Crossley D a. (2000) 14C distribution in soil organisms and respiration after the decomposition of crop residue in conventional tillage and no-till agroecosystems at Georgia Piedimont. *Soil Tillage Res.* 57: 31–41.

Van Groenigen KJ, Bloem J, Bååth E, Boeckx P, Rousk J, Bodé S, Forristal D & Jones MB (2010) Abundance, production and stabilization of microbial biomass under conventional and reduced tillage. *Soil Biol. Biochem.* 42: 48–55.

Van Groenigen KJ, Hastings A, Forristal D, Roth B, Jones M & Smith P (2011) Soil C storage as affected by tillage and straw management: An assessment using field measurements and model predictions. *Agric. Ecosyst. Environ.* 140: 218–225.

Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280: 309–316.

Izquierdo J a., Sizova M V. & Lynd LR (2010) Diversity of bacteria and glycosyl hydrolase family 48 genes in cellulolytic consortia enriched from thermophilic biocompost. *Appl. Environ. Microbiol.* 76: 3545–3553.

Joergensen RG (1996) The fumigation-extraction method to estimate soil microbial biomass: calibration of the kEC-value. *Soil Biol. Biochem.* 28: 25–31.

Kandeler E, Tscherko D & Spiegel H (1999) Long-term monitoring of microbial biomass, N mineralisation and enzyme activities of a chernozem under different tillage management. *Biol. Fertil. Soils* 28: 343–351.

Kellner H, Zak DR & Vandenbol M (2010) Fungi unearthed: transcripts encoding lignocellulolytic and chitinolytic enzymes in forest soil. *PLoS One* 5: e10971.

Küstermann B, Munch JC & Hülsbergen K-J (2013) Effects of soil tillage and fertilization on resource efficiency and greenhouse gas emissions in a long-term field experiment in Southern Germany. *Eur. J. Agron.* 49: 61–73.

Lal R (2004) Agricultural activities and the global carbon cycle. *Nutrient Cycling in Agroecosystems*, Vol. 70, pp. 103–116.

Langston JA, Shaghasi T, Abbate E, Xu F, Vlasenko E & Sweeney MD (2011) Oxidoreductive Cellulose Depolymerization by the Enzymes Cellobiose Dehydrogenase and Glycoside Hydrolase 61. *Appl. Environ. Microbiol.* 77: 7007–7015.

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM & Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42.

Lynd LR, Weimer PJ, van Zyl WH & Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* 66: 506–577, table of contents.

Meyer F et al. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9: 386.

Meyer-Aurich A, Gandorfer M, Gerl G & Kainz M (2009) Tillage and Fertilizer Effects on Yield, Profitability, and Risk in a Corn-Wheat-Potato-Wheat Rotation. *Agron. J.* 106: 1538–1547.

Miller RM & Lodge DJ (1997) 5 Fungal Responses to Disturbance□: Agriculture and Forestry. *The Mycota, Vol. Iv. Environmental and Microbial Relationships*, pp. 65–84.

Murugan R, Koch H-J & Joergensen RG (2013) Long-term influence of different tillage intensities on soil microbial biomass, residues and community structure at different depths. *Biol. Fertil. Soils.*

Ng JP, Hollister EB, González-Chávez MDC a., Hons FM, Zuberer D a., Aitkenhead-Peterson J a., Loeppert R & Gentry TJ (2012) Impacts of Cropping Systems and Long-Term Tillage on Soil Microbial Population Levels and Community Composition in Dryland Agricultural Setting. *ISRN Ecol.* 2012: 1–11.

Nyyssönen M, Tran HM, Karaoz U, Weihe C, Hadi MZ, Martiny JBH, Martiny AC & Brodie EL (2013) Coupled high-throughput functional screening and next generation sequencing for identification of plant polymer decomposing enzymes in metagenomic libraries. *Front. Microbiol.* 4: 282.

Oksanen J et al. (2013) vegan: Community Ecology Package. R package version 2.0-10. *R Packag. version* 1: 10.4135/9781412971874.n145.

Olson DG et al. (2010) Deletion of the Cel48S cellulase from Clostridium thermocellum. *Proc. Natl. Acad. Sci. U. S. A.* 107: 17727–17732.

Pritsch K, Luedemann G, Matyssek R, Hartmann A, Schloter M, Scherb H & Grams TEE (2005) Mycorrhizosphere responsiveness to atmospheric ozone and inoculation with Phytophthora citricola in a phytotron experiment with spruce/beech mixed cultures. *Plant Biol.* 7: 718–727.

Punta M et al. (2012) The Pfam protein families database. *Nucleic Acids Res.* 40: D290–D301.

R Core Team (2013) *R: A Language and Environment for Statistical Computing.* Vienna.

Reichenbach H, Lang E, Schumann P & Spröer C (2006) Byssovorax cruenta gen. nov., sp. nov., nom. rev., a cellulose-degrading myxobacterium: rediscovery of “Myxococcus cruentus” Thaxter 1897. *Int. J. Syst. Evol. Microbiol.* 56: 2357–2363.

Reichenbecher M, Lottspeich F & Bronnenmeier K (1997) Purification and properties of a cellobiose phosphorylase (CepA) and a cellodextrin phosphorylase (CepB) from the cellulolytic thermophile Clostridium stercorarium. *Eur. J. Biochem.* 247: 262–267.

Rho M, Tang H & Ye Y (2010) FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res.* 38: e191.

Schimel DS (1995) Terrestrial ecosystems and the carbon cycle. *Glob. Chang. Biol.* 1: 77–91.

Schlesinger WH & Andrews JA (2000) Soil respiration and the global carbon cycle. *Biogeochemistry* 48: 7–20.

Silva AP, Babujia LC, Matsumoto LS, Guimarães MF & Hungria M (2013) Bacterial diversity under different tillage and crop rotation systems in an oxisol of Southern Brazil. *Open Agric. J.* 7: 40–47.

Solans M & Vobis G (2003) Saprophytic actinomycetes associated to the rhizosphere and rhizoplane of Discaria trinervis. *Ecol. Austral* 13: 97–107.

Souza RC, Cantão ME, Vasconcelos ATR, Nogueira MA & Hungria M (2013) Soil metagenomics reveals differences under conventional and no-tillage with crop rotation or succession. *Appl. Soil Ecol.* 72: 49–61.

Tebrügge F, Düring R -a., Du R & Tebru F (1999) Reducing tillage intensity - A review of results from a long-term study in Germany. *Soil Tillage Res.* 53: 15–28.

Varner JE & Louis S (1989) Plant Cell Wall Architecture. *Cell* 56: 231–239.

Větrovský T, Steffen KT & Baldrian P (2014) Potential of cometabolic transformation of polysaccharides and lignin in lignocellulose by soil Actinobacteria. *PLoS One* 9.

White TJ, Bruns S, Lee S & Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols: A Guide to Methods and Applications*, pp. 315–322.

Woo HL, Hazen TC, Simmons B a & Deangelis KM (2013) Enzyme activities of aerobic lignocellulolytic bacteria isolated from wet tropical forest soils. *Syst. Appl. Microbiol.* 1–8.

Wooley JC, Godzik A & Friedberg I (2010) A primer on metagenomics. *PLoS Comput. Biol.* 6: e1000667.

Yokota A, Mathematics F, Sciences N, Indonesia U, Bakteri K, Chloroflexi F & Strain-strain KT (2012) Cultivation of Uncultured Bacteria of the Class Ktedonobacteria in the Phylum Chloroflexi. *Makara J. Sci.* 1: 1–8.

Figure Legends

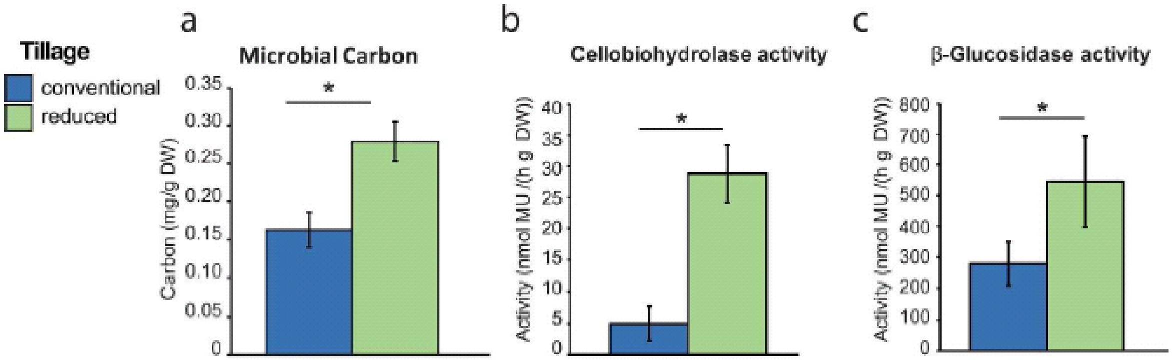

Figure 1: Reduced tillage increases microbial carbon and potential cellobiohydrolase and β-glucosidase activities in the topsoil: a) Depicted is the amount of microbial carbon detected in soils with conventional and reduced tillage treatments, in mg per gram dry weight soil. b,c) Shown is the amount of produced methylumbelliferone (MU) in nmol per hour and gram dry weight soil, that has been released by the enzymatic degradation of MU-β-cellobioside (b) or MU-β-D-glucoside (c). Significant differences between conventional and reduced tillage were determined by unpaired t-test statistics (* = *P*<0.05, n=3).

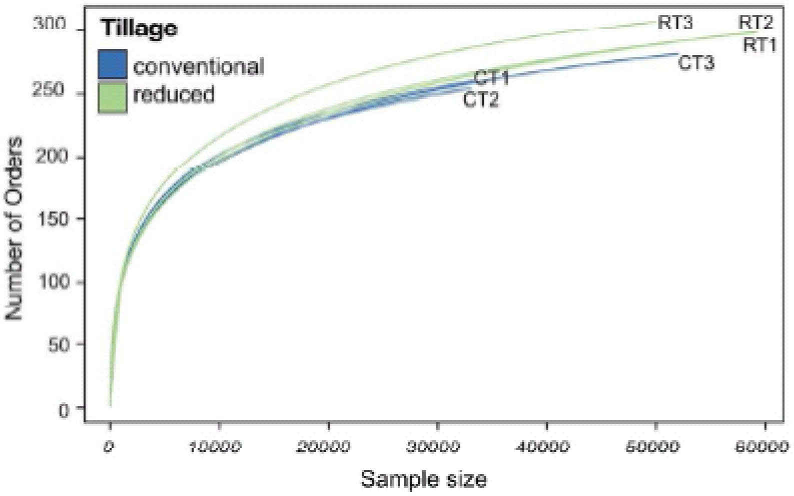

Figure 2: Rarefaction curves for the metagenomes of conventional and reduced tillage-treated soil: Depicted is the number of detected orders (Blastx (Diamond) of all filtered metagenome reads against the NCBI non-redundant protein database, e-value cut-off = 10^{-4}, assignment of taxonomic rank with MEGAN5) as a function of the sequencing depth, i.e. a randomly chosen amount of reads from the metagenomes for each replicate.

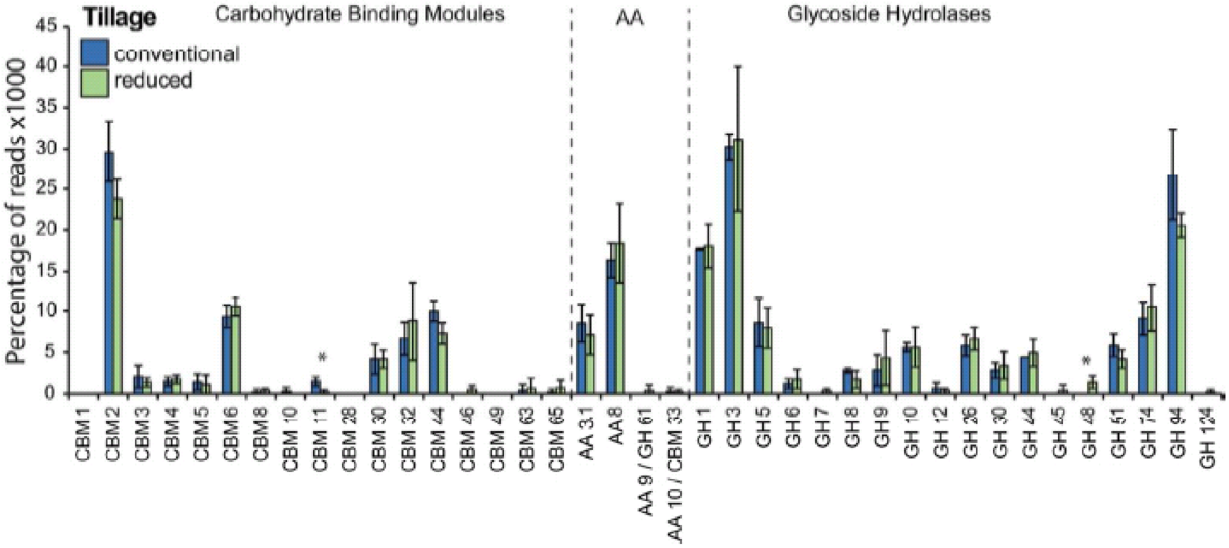

Figure 3: Prediction of cellulases and carbohydrate binding modules from metagenomes of conventional and reduced tillage-treated soil: Depicted is the percentage of reads multiplied by 1000 annotated to catalytic and carbohydrate binding modules found in the metagenomes of conventional and reduced tillage-treated soil using HMM modules and after an additional quality control using a curated database. Significant differences between conventional and reduced tillage were determined by unpaired t-test statistics (* <0.05, n=3).

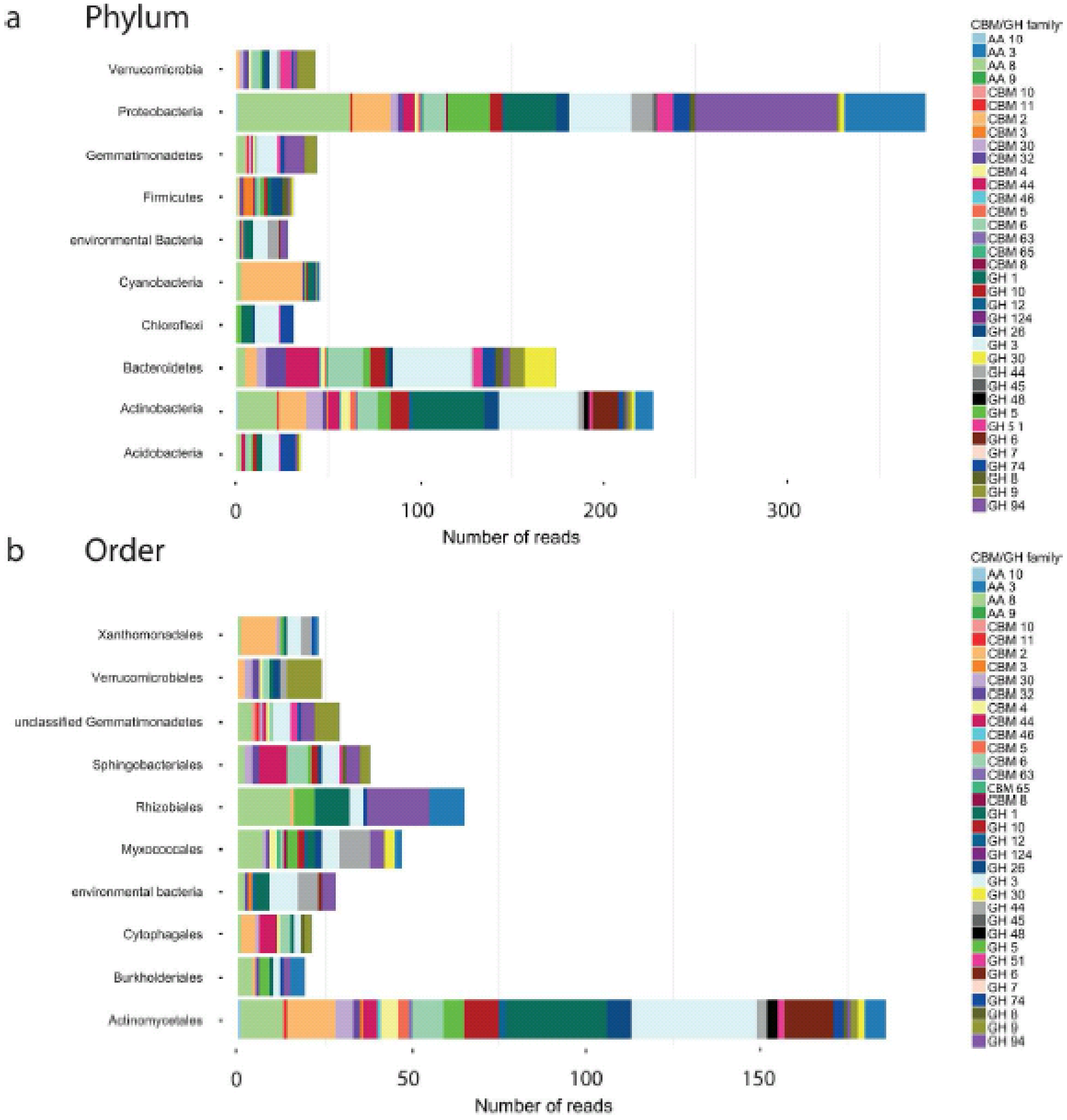

Figure 4:Taxonomic analysis of predicted catalytic and carbohydrate binding modules: Shown is the number of reads annotated to different phyla **a)** or orders **b)** (Blastx (Diamond) against NCBI non-redundant protein database, assignment of taxonomic ranks using MEGAN5) for each catalytic or carbohydrate binding module. Presented are the most abundant taxa.

Supplementary Figure Captions

Figure S1: Taxonomic analysis of metagenomes of conventional and reduced tillage-treated soil: Depicted are the percentage of reads of the thirty most abundant phyla (a) and orders (b), according to the NCBI non-redundant protein database (n=3).Taxonomic ranks were assigned using MEGAN5.

Figure S2: Principal component analysis of metagenomes of conventional and reduced tillage- treated soil: Depicted are the first two components of the principle component analysis based on the relative abundance of orders identified by Blastx (Diamond) against the NCBI non-redundant protein database and assignment of taxonomic ranks using MEGAN5. There is no clear clustering of the metagenomes indicating no strong differences between the metagenomes at this global level.

Figure S3: Functional analysis of metagenomes of conventional and reduced tillage-treated soil: Shown is the percentage of reads annotated (Blastx (Diamond) against the KEGG database, e-value cut-off=10^{-5} and assignment of functionality using MEGAN5) to different KEGG-pathways, from metagenomes of conventional and reduced tillage-treated soil (n=3). Presented are the thirty most abundant pathways.

Figure S4: Taxonomic analysis of most abundant catalytic and carbohydrate binding modules: For the six most abundant HMM families from Figure 3, a detailed taxonomic annotation with the most abundant phyla is shown. Significant differences between conventional and reduced tillage were determined by unpaired t-test statistics.