Effects of dry- and wet-sieving of soil on identification and interpretation of microbial community composition

Aimeric Blaud[a, 1]*, *Manoj Menon*,[a, 2], *Bas van der Zaan*[b], *Georg J. Lair*[c, d], *Steve Banwart*[a, 3]

[a] *Department of Civil and Structural Engineering, Kroto Research Institute, The University of Sheffield, Broad Lane, Sheffield S3 7HQ, United Kingdom.*

[b] *Deltares, Subsurface and Groundwater Systems, Princetonlaan 6-8, 3508 Al Utrecht, the Netherlands.*

[c] *University of Natural Resources and Life Sciences (BOKU), Institute of Soil Research, Vienna, Peter-Jordan-Str. 82, 1190 Vienna, Austria.*

[d] *University of Innsbruck, Institute of Ecology, Sternwartestr. 15, 6020 Innsbruck, Austria.*

*Corresponding Author.

E-mail address: aimeric.blaud@gmail.com

[1] Current address: Agroecology Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK.

[2] Current address: Department of Geography, The University of Sheffield, Sheffield, S10 2TN, UK.

[3] Current address: School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK.

This is the pre-print version of the article. Please refer to the following link for the printed version: http://www.sciencedirect.com/science/article/pii/S0065211316301122

Published in /how to cite this article/chapter:

A. Blaud, M. Menon, B. van der Zaan, G.J. Lair and S.A. Banwart, Effects of Dry and Wet Sieving of Soil on Identification and Interpretation of Microbial Community Composition. In: Steven A. Banwart and Donald L. Sparks, editors, Advances in Agronomy, Vol. 142, Burlington: Academic Press, 2017, pp. 119-142.

Abstract

Soil aggregates are microhabitats for microorganisms, and directly influence microorganisms that live within and are influenced by microorganisms in return. Two methods are used to isolate soil aggregates by their size: dry- (sieving air dried soil) or wet-sieving (sieving soil in water). Wet-sieving methods are generally considered to represent separation of aggregate classes that are stable to physical dis-aggregation in water, a condition considered favourable for protecting soil structure over time. However, little is known about the effect of sieving methods on microbial abundance, diversity and functions, hindering the understanding of the relationship between soil structure and soil aggregates as habitat and soil microorganisms. In this study, the effect of dry- and wet-sieving on bacterial diversity, and abundance of microorganisms involved in N fixation (*nifH* gene), nitrification (*amoA* bacteria and archaea) and denitrification (*narG, nirS* and *nosZ* genes), were determined for 4 sizes of soil aggregates from a cropland and grassland. Quantitative-PCR (Q-PCR) showed little differences in relative gene abundance between size fractions of soil aggregates, but wet-sieving method significantly increased gene abundance for *amoA* bacteria, *nirS* and *nosZ* genes. When the N functional genes were expressed as percentage of the bacterial 16S rRNA genes, the wet sieving resulted in significantly higher genes percentage for all the genes (except for *narG* gene), and significant differences between soil aggregate size fractions at the grassland site. The different sieving methods resulted in different bacterial community compositions, but only the wet-sieving method was able to reveal significant differences in bacterial community composition between soil fractions in grassland. The results demonstrate significantly different quantitative and qualitative interpretation of soil microbial community depending on whether aggregate samples were obtained from wet- or dry-sieving, highlighting the importance in the choice of the sieving method.

Keywords: Quantitative-PCR, Amplicon sequencing, nitrogen fixation, nitrification, denitrification, soil aggregates, grassland, cropland

1. Introduction

Soil is an extremely complex and heterogeneous environment, due to the complexity of its structure (i.e. 3-D architecture of pores and particles), the large vertical spatial heterogeneity across the different horizons of soil profiles, and a huge and largely unknown microbial genetic diversity. Soil aggregates, composed of soil mineral fragments, decaying biomass, gases, water and solutes, and living organisms bound together as porous particles, represent the complexity of the soil structure and also the microhabitats for the microorganisms. Both the soil structure and soil microorganisms are central soil features that determine many key functions such as soil water retention and transmission, C, N, P, K sequestration, and nutrient transformations that ultimately sustain soil fertility. Different sizes of soil aggregates was shown to harbour different bacterial community structure, (Blaud et al., 2012; Fall et al., 2004; Helgason et al., 2010; Kandeler et al., 2000; Sessitsch et al., 2001; Vaisanen et al., 2005), different bacterial diversity (Davinic et al., 2012; Kravchenko et al., 2014; Sessitsch et al., 2001), bacterial abundance and biomass (Helgason et al., 2010; Mendes et al., 1999; Sainju, 2006; Schutter and Dick, 2002) and microbial activity (Bach and Hofmockel, 2014; Lensi et al., 1995; Sey et al., 2008). These differences are linked to the specific environmental conditions which exert biological selection pressures and are highly variable within aggregates.

To study soil aggregates, sieving methods are used to isolate different size classes of soil aggregates. The separation of soil aggregates is mainly done by dry- or wet- sieving methods. The wet-sieving method, first described by Yoder (1936), is the most commonly used method to study microbial communities in soil aggregates and involves immersing soil for several minutes in water to break down aggregates. This occurs by increasing the surrounding static water pressure on the air trapped inside immersed particle pores, followed by vertical strokes in water to create shear forces to separate the soil particles that are initially placed on the top of a nest of subsequently immersed sieves. Dry-sieving involves shaking usually air-dried soil, on top of a nest of sieves. Thus, the energy applied to the soil differs greatly between dry- and wet-sieving which affects directly the amount of stable soil aggregates that are obtained. Furthermore, wet-sieving affects the aqueous colloidal forces at particle surfaces that can enhance or diminish the cohesive forces between aggregated particles. Thus, these two methods are expected to have direct effect on microbial communities due to the different sizes of soil aggregates which are isolated, i.e. the

"washing" effect during wet sieving coupled with potential cross contamination between soil fractions, the effect of drying soil before dry-sieving, and the different mechanical and physical-chemical forces applied on soil aggregates.

Only few studies have investigated the impact of dry- and wet-sieving to separate soil aggregates. Most of these studies focused on the effects of sieving methods on the physico-chemical characteristics of soil aggregates. Dry-sieving maintains large soil aggregates sizes (> 2 mm) but is usually limited to the size fractions > 250 µm. In contrast, wet-sieving can separate soil aggregates from various size classes and in particular smaller sizes (< 250 µm). The proportion of soil aggregates with size < 2 mm mainly increase with wet-sieving while soil aggregates > 2 mm decrease due to the breakdown of the macroaggregates into smaller aggregates, and inversely for dry-sieving (Beauchamp and Seech, 1990; Sainju, 2006; Bach and Hofmockel, 2014). Wet-sieving leads to a loss of total C or total N, especially for soil fractions < 250 µm, although no change or sometimes an increase in C content (for either > 250 and < 250 µm soil fractions) were found for wet-sieving in comparison to dry-sieving (Sainju, 2006). Seech and Beauchamp (1988) concluded that wet-sieving methods result in underestimating C and N pools.

The impact of aggregate fractionation procedures on microbial communities is not well studied. Sainju (2006) showed that the wet-sieving method decreases the nitrogen microbial biomass in comparison to dry-sieving. In contrast, the carbon microbial biomass can decrease or increase depending on the soil type (Sainju, 2006). However the C or N microbial biomass is a gross indicator of microbial biomass, and no study has investigated the effect of sieving methods on microbial abundance, community structure or diversity using DNA-based approaches (e.g. Q-PCR, next generation sequencing). A recent study comparing the effect of dry- and wet-sieving on microbial enzymatic activity showed that wet-sieving overestimated the potential microbial enzymatic activity in comparison to dry-sieving (Bach and Hofmockel, 2014). However, only the enzymatic activity differed between sizes of soil aggregates with wet-sieving and not with dry sieving. This study also showed that drying the samples at 4 °C to reach 10-20% of soil gravimetric water content did not affect the enzymatic activities before dry-sieving.

The effect of sieving methods on microbial communities and resulting microbial characterisation data and their interpretation remains largely unknown. This gap in understanding may represent a major factor influencing the results of any study investigating microbial communities in soil aggregates, and is limiting the understanding of the

relationship between soil structure, soil functions and microbial diversity. Thus, the objective of this study was to determine the effect of dry- and wet-sieving on microbial community abundance and diversity within different size classes of soil aggregates. Four sizes of soil aggregates from a cropland and grassland were obtained by dry- and wet-sieving. Then, the abundance of bacteria, fungi and microbial communities involved in N fixation, nitrification and denitrification, and bacterial diversity were determined by quantitative PCR and amplicon sequencing respectively, for each size class of soil aggregates and for the bulk soil.

2. Material and methods

2.1 Study sites and soil sampling

The study sites, a cropland and grassland are located east of Vienna, Austria, in the National Park ‘‘Donau-Auen’’ on a floodplain of the Danube River. The cropland site was a grassland since 1781 and was converted to intensive cropland in the first half of the 20th century. The grassland site was converted from forest to grassland (presently *Onobrychido viciifoliae-Brometum*) between 1809 and 1859 and is currently cut twice a year. The topsoil (0-10 cm) age is approx. 250-350 years since deposition of fluvial sediments as parent material forming a terrace above the down cutting river channel (Lair et al., 2009). The soils are classified as Mollic Fluvisols (IUSS Working Group WRB, 2014). The soil characteristics for cropland and grassland are shown in Table 1. Both sites were sampled on the 27th of September 2013. Three distinct soil samples (500 g) were sampled at each site (grassland and cropland sites) from 5-10 cm depth and store at 4 °C until soil fractionation.

2.2 Soil fractionation

The soil samples were sieved at 2 mm before dry- or wet-sieving to homogenise the samples and to remove large roots and stones. Dry- and wet- sieving were performed on all the replicate samples for each site. Twenty grams of soil were used for each soil fractionation by size, recovered from the sieves of specific screen sizes. Henceforth, the term “soil fraction” is preferred to “soil aggregates” because this study did not separate soil aggregates from single mineral particles.

Table 1. Soil characteristics of bulk soil samples on a dry mass basis. Mean value ± one standard deviation (n = 3) are shown.

Location	Cropland 48°09'N, 16°41'E	Grassland 48°11'N, 16°44'E
Water content (%)	22.0 ± 2.9	15.8 ± 2.0
Soil pH (H_2O)	7.7 ± 0.14	7.4 ± 0.09
Organic C (%)	2.4 ± 0.36	5.0 ± 0.60
Total N (%)	0.13 ± 0.01	0.33 ± 0.04
C_{org}/N	18.1 ± 1.83	15.0 ± 0.52
$N\text{-}NH_4^+$ (mg kg^{-1})	1.59 ± 0.29	4.77 ± 0.98
$N\text{-}NO_3^-$ (mg kg^{-1})	20.3 ± 3.07	1.5 ± 0.66
$P\text{-}PO_4^{3-}$ (g kg^{-1})	0.35 ± 0.10	0.59 ± 0.04
$CaCO_3$ (%)	19.0 ± 1.90	21.1 ± 1.41
Sand, 63-2000 μm (%)	32.7	8.2
Silt, 2-63 μm (%)	43.8	63.0
Clay, < 2 μm (%)	23.5	28.8

2.2.1 Dry-sieving

Prior to dry-sieving, the 2 mm sieved soils were air-dried at 4 °C for 7 days until they reached a gravimetric water content of ~80 g kg^{-1} (Sainju et al., 2003). The air-drying was required to obtain the soil fraction < 53 μm from grassland soil and any soil fractions < 250 μm from cropland soil. The dry sieving protocol consisted of shaking by hand the soil samples placed on top of a nest of sieves (1000, 250 and 53 μm; 10 cm Ø) for 3 min at ~200 rotation min^{-1} (Sainju et al., 2003; Sainju, 2006). Soil retained on the 1000, 250 and 53 μm sieves were considered as 1000-2000 μm, 250-1000 μm and 53-250 μm soil fractions, respectively. The soil collected in the cup under the 53 μm sieve was the < 53 μm soil fraction. Soil aliquots were taken directly from each sieve for DNA extraction and stored at -20 °C, and the rest of the soil fractions were dried at 55 °C and used to measure soil fractions' mass distributions.

2.2.2 Wet-sieving

The wet-sieving fractionation method was adapted from Yoder (1936) and Blaud et al., (2012). Fresh soil samples were place on top of a nest of sieves (1000, 250 and 53 µm; 10 cm Ø) and immersed in ~1.3 l ultra-pure sterile water (4 °C) tank for 5 min. Then, the sieves were raised and lowered during 10 min (stroke length ~30 mm, frequency 30 cycles min^{-1}). Soil retained on the 1000, 250 and 53 µm sieves were considered as 1000-2000 µm, 250-1000 µm and 53-250 µm soil fractions, respectively. The water and soil left in the tank were centrifuged at 4500 G for 10 min. The centrifugation was repeated to reduce the volume of water as much as possible and collect the soil particles, which represented the soil fraction < 53 µm. Two soil aliquots were taken directly in each sieve: one for DNA extraction placed at -20 °C, and one for soil water content measurement dried at 55 °C. The rest of the soil fractions for each sieve were washed in tubes and dried at 55 °C and used to measure soil fractions' mass distributions.

The pore liquid collected after each round of centrifugation was filtered at 0.22 µm (47 mm Ø GTTP filter, Wathman) in order to collect and quantify the microorganisms washed from soil fractions during the sieving method. For each sample, 5 filters were required to filter the entire volume of water (due to clogging of the filter), except for two replicates of cropland that required 6 and 7 filters. The filters were kept at -20 °C before DNA extraction.

2.3 DNA extraction

DNA was extracted from 0.25 g of fresh soil for each soil fraction and bulk soil (i.e. 2 mm sieved soil) using the PowerSoil® DNA Isolation Kit (Mo-Bio laboratories, Carlsbad, CA, USA) according to manufacturer's instruction, except for the final step where the nucleic acids were eluted in 100 µl of sterile nuclease-free water.

DNA was extracted from the water used for wet-sieving (after centrifugation to obtain < 53 µm soil fraction) to determine the relative abundance of microorganisms lost during wet sieving. The same amount of water without soil was also filtered and used as control to ensure that the result obtained came from the wet-sieving and not from contamination of the water or filter. The water for each sample was filtered and DNA was extracted from the filter using the PowerWater® DNA isolation kit (Mo-Bio laboratories, Carlsbad, CA, USA) according to manufacturer's instruction, except for the final step where the nucleic acids were

eluted in 100 µl of sterile nuclease free water. DNA was extracted for each filter (i.e. 33 filters in total) and the DNA extracts were pooled for each sample.

2.4 Quantitative-PCR

Variation in microbial gene abundance was determined by Quantitative-PCR (Q-PCR) targeting specific genes or genetic regions. Bacterial community was targeted via the 16S rRNA gene while the fungal community abundance was investigated by targeting the ITS region. The different communities involved in most steps of the N-cycle were investigated: the nitrogen fixing microorganisms were quantified based on the *nifH* gene; nitrification was investigated by targeting the ammonia oxidising bacteria (AOB) and archaea (AOA) via the *amoA* gene, and denitrifiers were targeted via the *narG* gene coding for the nitrate reductase, the *nirS* gene coding for the nitrite reductase and the *nosZ* gene coding for the nitrous oxide reductase. The details of the primers used to amplify the different amplicons are given in Table S1.

Q-PCR standards for each molecular target were obtained using a 10-fold serial dilution of plasmids carrying a single cloned target gene, constructed by cloning PCR product of environmental samples (pCR2.1 TOPO vector), isolating cloned inserts (Qiagen Plasmid mini Kit), and checking for the presence of gene of interest by sequence-analysis. Standard curves and the no template control were amplified in triplicate in the same plate as the environmental samples. Q-PCR amplifications were performed in 25 µl volumes containing 12.5 µl of iQ™ SYBR® Green Supermix (Bio-Rad, Hemel Hempstead, UK), 8.5 µl of nuclease-free water (Ambion, Warrington, UK), 1.25 µl of each primer (10 µM) and 1 µl of template DNA using a CFX96™ Real-Time System (Bio-Rad, Hemel Hempstead, UK). Standard amplification was used for all Q-PCR assays except AOA, starting with an initial denaturation at 95 °C for 3 min, followed by 40 cycles of 30 s at 95 °C, 0.5 to 1 min of annealing (annealing temperature and time for each primers pairs are given in Table S1), and 30 s at 72 °C (Tsiknia et al., 2013). The fluorescence was measured at the end of each synthesis step (i.e. at 81 °C for AOA and at 72 °C for all other genes).

Threshold cycle (Ct) values and amplicon numbers were determined automatically using the Bio-Rad CFX Manager™ software. The efficiency of the Q-PCR assays varied between 70-98%. The r^2 were > 0.99 for all the genes, except for *nifH* gene (0.984). The presence of Q-PCR inhibitors was tested for bacterial 16S rRNA gene, by running a Q-PCR

with DNA extracts 10 times diluted or mixed with a known amount of the standard. No inhibition was detected.

The specificity of the Q-PCR was assessed via a melting curve analysis (increase of temperature from annealing temperature to 95 °C by 0.5 °C per step of 0.05 s) at the end of each Q-PCR amplification (Ririe et al., 1997). The melting curves for the bacterial 16S rRNA, *nifH*, *amoA*, *narG*, *nirS*, and *nosZ* genes Q-PCR assays showed specificity for the amplified targeted genes (i.e. single peak). As expected, the melting curve of the Q-PCR for fungal ITS showed the amplification of products of different lengths, due to the variability in length of ITS regions between different fungal taxa (Manter and Vivanco, 2007).

2.5 Amplicon sequencing

The bacterial diversity of the different soil fractions obtained by dry- and wet-sieving, bulk soil and microbial suspension from water of the wet-sieving, for the cropland and grassland was determined using the Ion Torrent® platform. The bacterial 16S rRNA gene V4 variable region was amplified using the primers 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GGACTACHVG GGTWTCTAAT-3′) (Caporaso et al., 2011) in a single-step 30 cycle PCR using the HotStarTaq Plus Master Mix Kit (Qiagen, USA) and the following conditions: 94°C for 3 min, followed by 28 cycles (5 cycles used on PCR products) of 94°C for 30 s, 53°C for 40 s and 72°C for 1 min, followed by a final elongation step at 72°C for 5 min. Amplicon sequencing was performed at MR DNA (www.mrdnalab.com, Shallowater, TX, USA) on an Ion Torrent PGM following the manufacturer's guidelines.

The PGM data were analysed following the pipeline developed by Pylro et al (2014) that uses UPARSE (Edgar, 2013) and QIIME (Caporaso et al., 2010). Briefly, strip barcode, quality filtering, dereplication, abundance sort and discard singletons were done using USEARCH 1.8. Chimera filtering was done using the *rdp_gold.fa* dataset. Then, taxonomy was assigned to operational taxonomic unit (OTU) using *uclust* method on QIIME 1.8 and Greengenes data base (13_8) as a reference. The number of bacterial sequences per sample was on average 9183 ± 1443. Few archaeal sequences were found with on average 174 ± 99 per sample.

2.6 Statistical analysis

To determine differences in the means of aggregate distribution, microbial gene abundance or bacterial phylum relative abundance, ANOVA tests were performed with sites,

sieving methods and soil fractions as factors. The normality of the model residuals and the homoscedasticity of the variances were checked before statistical analysis. Log transformations of the Q-PCR data were applied to meet these criteria, except for *narG* gene abundance. When significant differences were found by ANOVA, the post-hoc test of Newman-Keuls was performed to reveal the significance differences between class pairs. To test the differences between sites of the loss of genes in the wet-sieving water, the Student test was used.

The bacterial community composition was visualised by Principal Coordinate Analysis (PCoA) based on the relative abundance of the OTU and generated using Bray-Curtis distance. ANOSIM (Analysis of SIMilarity; 10,000 maximum permutations) was used to investigate potential differences between bacterial community composition due to sieving-method, site or soil fractions (Clarke and Green, 1988). Two-way ANOSIM was used to compare one factor against the other factors and one-way ANOSIM to investigate the influence of an individual factor. ANOSIM analysis yields an R value, whereby ANOSIM values close to R = 1 indicate a high separation between groups (e.g. between soil fractions), whilst ANOSIM values close to R = 0 indicate a low group separation.

ANOVA and PCoA were performed using R v3.2.1 (R Development Core Team, 2015) and the package Phyloseq for PCoA (McMurdie and Holmes, 2013), while the ANOSIM tests were performed using PRIMER software (v6, PRIMER-E Ltd, Plymouth, UK).

3. Results

3.1 Aggregates distribution

The soil fractionation procedure resulted in mean mass recovery ~100% of the original unfractionated soil. The aggregates mass distribution showed similar pattern within both sites. The soil fractions > 250 µm represented 35-50 % of the aggregate distribution, while the soil fractions < 250 µm were significantly lower and represented 2-20% (Fig 1; Table S2). In contrast, the aggregate distribution from the cropland obtained by wet sieving showed the opposite distribution compared to any other distribution, with an increase in the mass of soil fractions with decreasing size of soil factions. The soil fraction 1000-2000 µm in cropland was significantly lower than < 53 µm fraction (~16% and ~35% of the aggregate distribution, respectively). The mass distribution of each soil fraction for cropland was significantly different between sieving methods, except for the 250-1000 soil fraction. The

aggregate distribution for cropland obtained by wet-sieving showed large standard errors in comparison to any other aggregate distribution. The sieving methods had also some effect for the grassland, with an increase by ~10% of the 250-1000 µm fraction with dry-sieving, and a significant increase by ~10% of the < 53 µm fraction with wet-sieving.

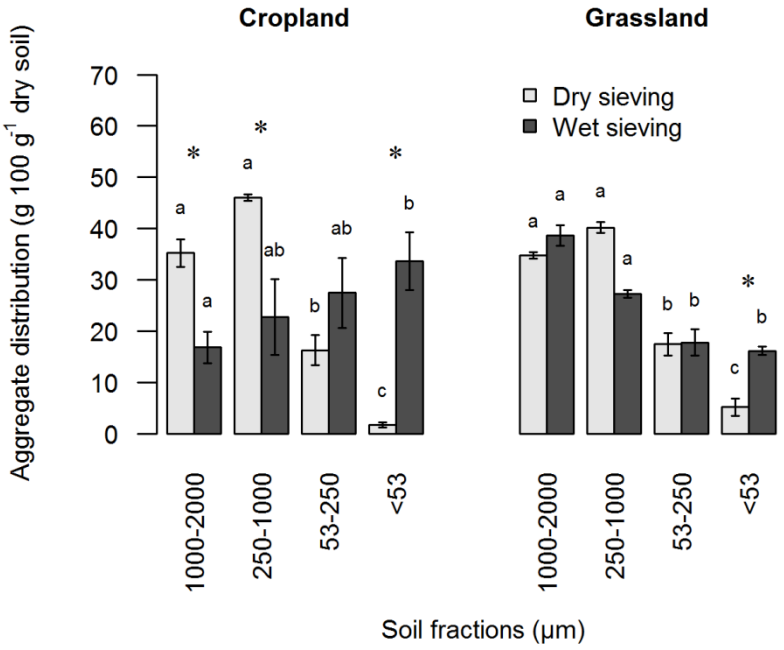

Fig. 1. Weight distribution of soil fractions (g 100 g^{-1} dry soil) obtained by dry- or wet-sieving method of soils from cropland and grassland. Means values ± standard error (n = 3) are shown. * indicates significant (P < 0.05) difference between dry- and wet-sieving for a specific soil fraction and site. Different letters indicate significant (P < 0.05) difference between soil fractions for a specific sieving method and site.

3.2 Microbial gene abundance

Microbial gene abundance showed significant differences (P < 0.01) between sites for all the genes except for *narG* gene (Fig. 2, Table S3). The genes abundance were higher in the grassland site for bacterial 16S rRNA gene, fungal ITS amplicon, *nifH*, *nirS* and *nosZ* genes. In contrast, *amoA* bacteria (AOB) gene showed higher abundance in cropland, while *amoA* archaea (AOA) showed slightly higher abundance in grassland. Only the bacterial 16S rRNA gene showed significant differences (P = 0.027; Table S3) between soil fractions, and the Post-hoc test revealed significant differences in grassland and dry-sieving between 1000-2000 µm and the fractions 250-1000 and 53-250 µm (Fig. 2). A significant effect (P < 0.001) of the sieving methods was found for the relative abundance of AOB, *nirS* and *nosZ*, with higher relative genes abundance found in fractions obtained by wet-sieving in grassland (Fig. 2; Table S3). However, the Post-hoc test did not reveal significant pair-wise differences.

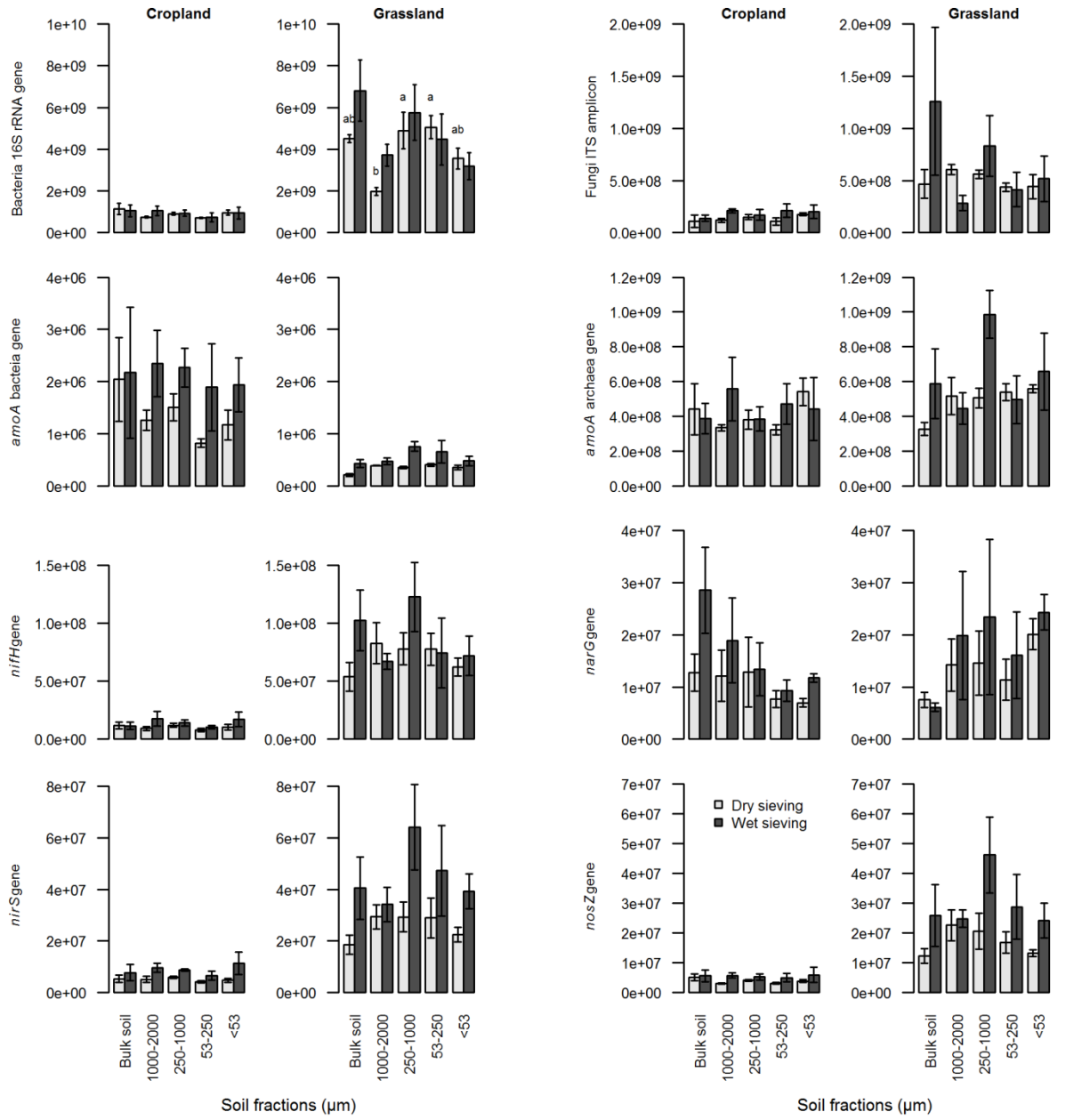

Fig. 2. Variation in gene abundance of bacteria (16S rRNA gene), fungi (ITS amplicon), N fixating (*nifH* gene), ammonia oxidizing bacteria and archaea (*amoA* gene), nitrate reductase (*narG* gene), nitrite reductase (*nirS* gene) and nitrous oxide reductase (*nosZ* gene) between four soil fractions obtained by dry- or wet-sieving methods from cropland and grassland. All abundances are expressed on the basis of 1 g of dry mass of soil fraction or bulk soil. Means values ± standard error (n = 3) are shown. * indicates significant (P < 0.05) different between dry- and wet-sieving for a specific soil fraction and site. Different letters indicate significant (P < 0.05) difference between soil fractions for a specific sieving method and site.

The proportion of microbial gene, expressed as percentage of bacterial 16S rRNA gene copies, was significantly (P < 0.001) different between sites, with higher *nifH* gene

proportion found in grassland than cropland, while higher proportions for AOB and *narG* genes were found in cropland (Fig. 3; Table S4). Significant difference between soil fractions and sieving methods were found for all the genes except for *narG* gene.

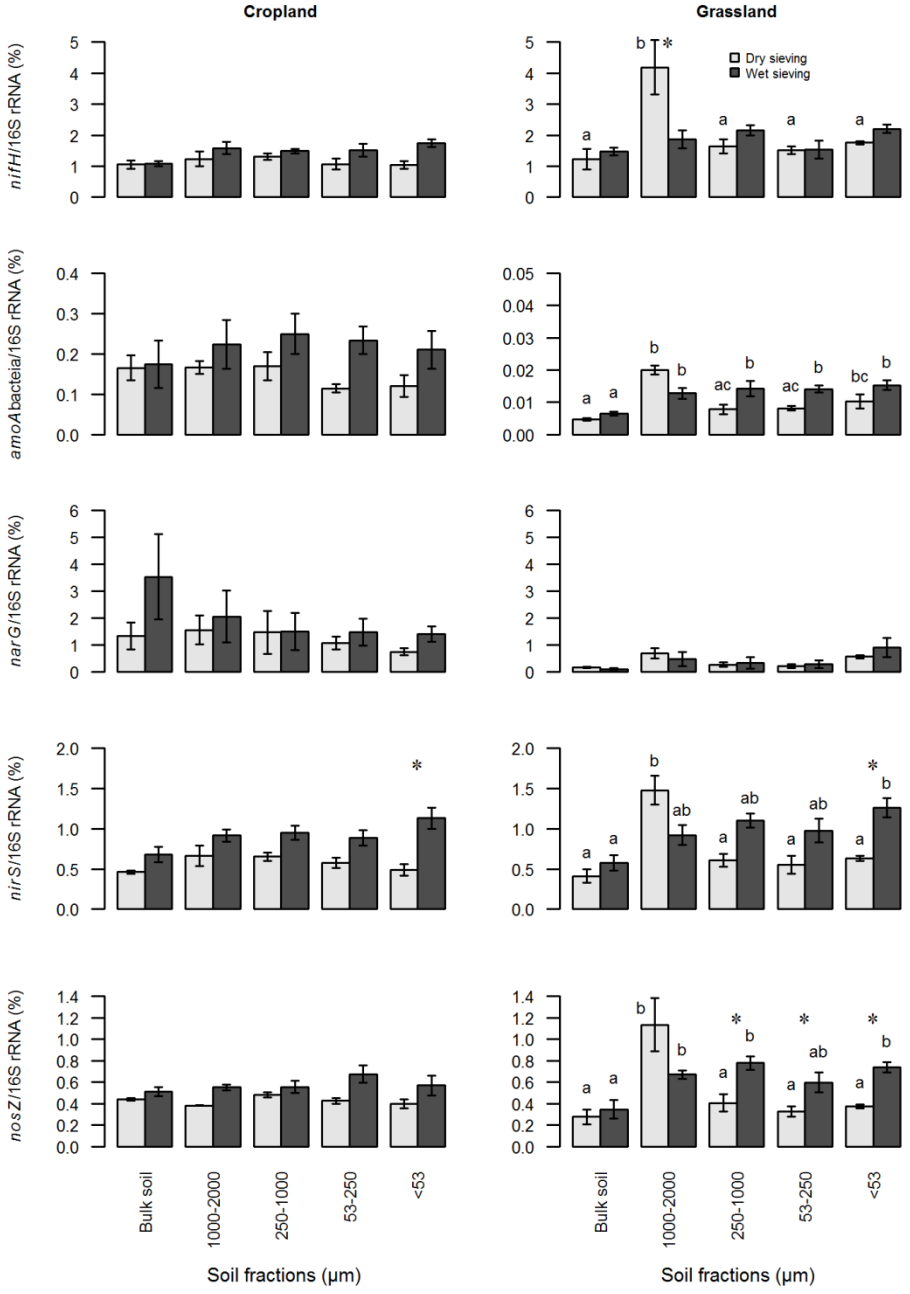

Fig. 3. Variation in N functional gene/bacterial 16S rRNA (%), of the N fixating (*nifH* gene), ammonia oxidizing bacteria (*amoA* gene), nitrate reductase (*narG* gene), nitrite reductase (*nirS* gene) and nitrous oxide reductase (*nosZ* gene) between four soil fractions obtained by dry- or wet-sieving methods from cropland and grassland. Means values ± standard error (n = 3) are shown. * indicates significant ($P < 0.05$) different between dry- and wet-sieving for a specific soil fraction and site. Different letters indicate significant ($P < 0.05$) difference between soil fractions for a specific sieving method and site.

The Post-hoc test revealed a similar trend between soil fractions for grassland obtained by dry-sieving, with the 1000-2000 µm fraction showing significantly (P < 0.05) higher proportion of microbial genes in comparison to most soil fractions and bulk soil (Fig. 3). The soil fractions from grassland obtained by wet-sieving showed higher proportions of AOB, *nirS* and *nosZ* genes than bulk soil, but no significant differences between soil fractions were found. The effect of sieving methods, showed higher genes proportions with wet-sieving by ~0.5%, except for the 1000-2000 µm fraction for grassland that showed higher proportion of *nifH*, AOB, *nirS* and *nosZ* gene with dry-sieving by 0.5% to 2%. The Post-hoc test revealed significant (P < 0.05) differences in gene proportions between sieving methods for *nifH*, *nirS* and *nosZ* genes for grassland, and *nirS* gene for cropland (Fig. 3).

The microbial genes abundance lost in the water during wet-sieving were expressed as percentage of the same gene present in 1 g of bulk soil. The proportion of microbial genes found in the sieving water varied between 0.3 to 2.3% (Table 2). Only *narG* gene showed ~7% of gene copies lost in sieving water for grassland, and was also the only gene with a significant (P = 0.0075) difference between sites. The microbial gene abundance in the sieving water was consistently higher in grassland than cropland and significant (P < 0.05) for bacteria, fungi, *nifH*, *narG* and *nosZ*, and marginally significant for AOA and AOB (P = 0.06 and 0.053, respectively; Fig. S1).

Table 2. Proportion of genes (%) lost in the water during soil fractionation using wet-sieving. The loss of gene number in the water is express as a percentage of the number of the same gene present in 1 g of bulk soil. Mean value ± one standard error (n = 3) are shown. Different letter indicate significant (P < 0.01) differences between cropland and grassland for a specific gene.

Gene	Cropland	Grassland
Bacterial 16s rRNA	1.55 ± 0.43	0.75 ± 0.30
Fungal ITS	0.48 ± 0.11	0.71 ± 0.52
nifH	2.31 ± 0.84	1.90 ± 0.85
amoA bacteria	0.33 ± 0.12	2.14 ± 0.63
amoA archaea	0.83 ± 0.09	1.83 ± 0.60
narG	1.16 ± 0.41 A	6.97 ± 0.80 B
nirS	0.85 ± 0.31	0.57 ± 0.17
nosZ	0.45 ± 0.14	0.60 ± 0.18

3.3 Bacterial diversity

The PCoA showed that the bacterial community composition from the water of wet-sieving differed greatly in comparison to any other samples (Fig. 4). The PCoA and ANOSIM also showed that the bacterial composition differed significantly (R = 0.45, *P* = 0.0001) between dry- and wet- sieving although some samples were mixed within each group. Then, significant differences between cropland and grassland were found, showing similar ANOSIM values compared to those reflecting the effect of the sieving methods (R = 0.45, *P* = 0.0007). The ANOSIM also revealed significant differences between soil fractions, bulk soil and water fractions but with a lower R value than those obtained for sieving methods and sites (R = 0.32, *P* = 0.0001).

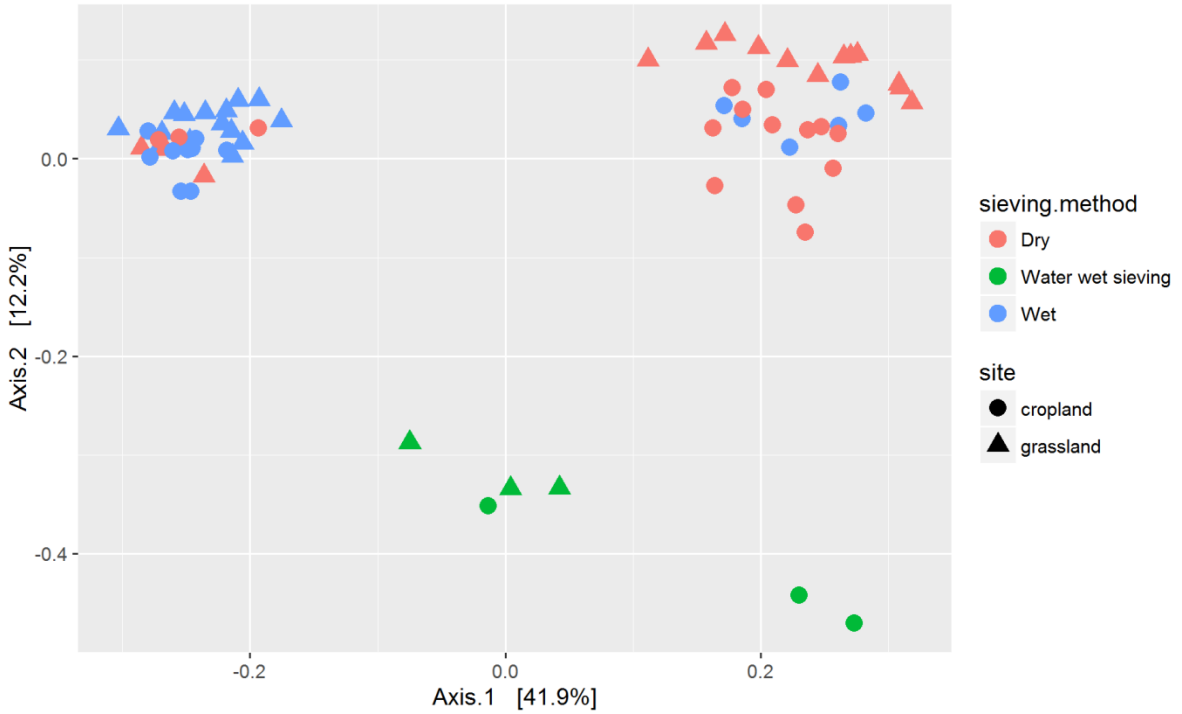

Fig. 4 PCoA of bacterial community of four soil fractions obtained by dry- or wet-sieving method and bulk soil from cropland and grassland. The PCoA was based on relative abundance of OTU and generated using Bray-Curtis distance. The six samples (in green) isolated from the rest of the samples correspond to water from the wet-sieving.

The PCoA and ANOSIM were also performed on soil fractions and bulk soil for each site to reveal how the sieving methods affected the bacterial community composition between soil fractions at each site, which was not visible on the global analysis (Fig. 5). Significant differences between sieving methods and between soil fractions were found for grassland (sieving: R = 0.82, *P* = 0.0001; fractions: R = 0.56, *P* = 0.0001) but not for cropland (*P* >

0.2). The ANOSIM analysis performed on grassland for each sieving-method revealed significant differences between soil fractions or bulk soil with both sieving methods (dry-sieving: R = 0.57, *P* = 0.0001; fractions: R = 0.58, *P* = 0.0001). The bulk soil showed clear

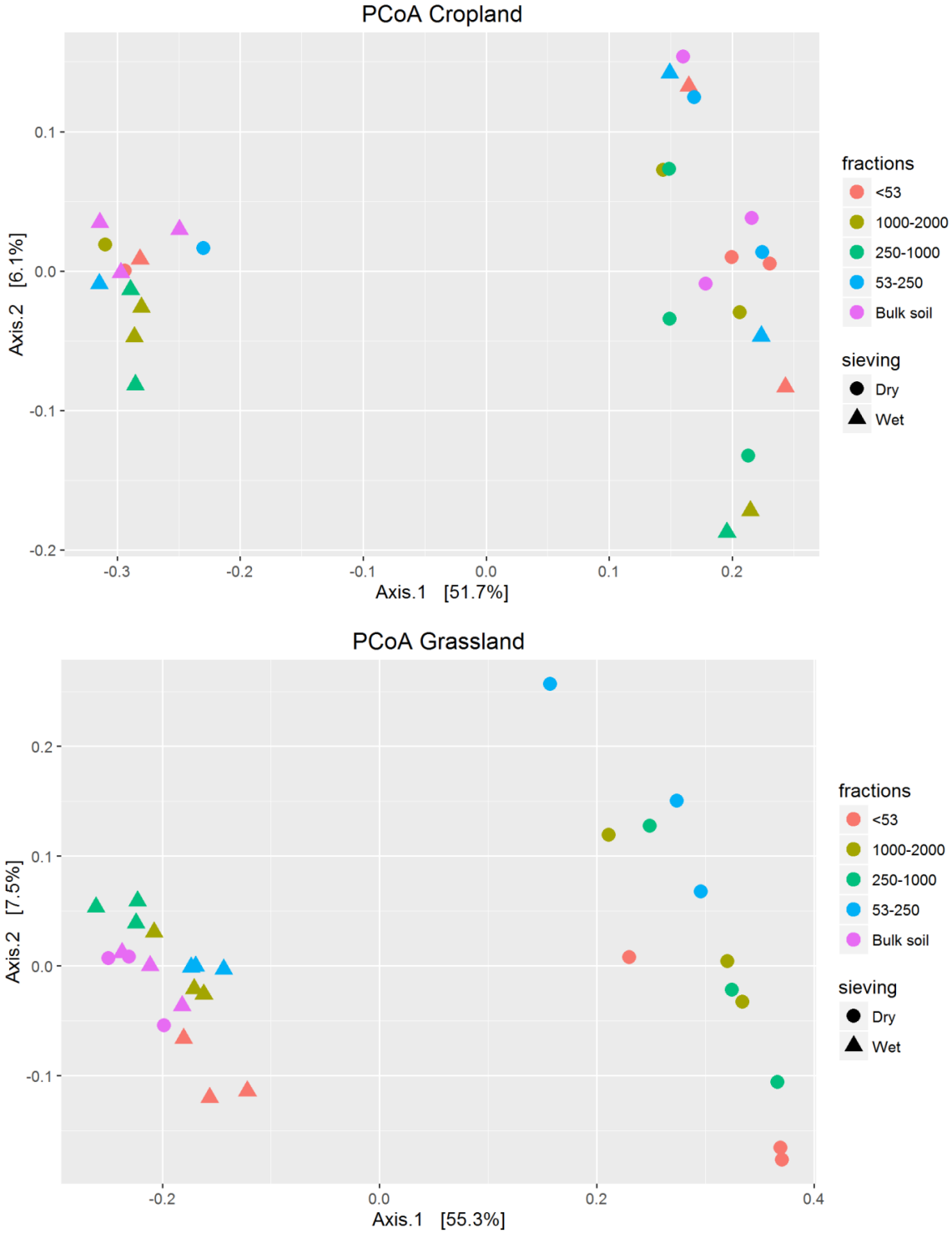

Fig. 5 PCoA of bacterial community of four soil fractions obtained by dry- or wet-sieving method and bulk soil from cropland (top) and grassland (bottom). The PCoA were based on relative abundance of OTU and generated using Bray-Curtis distance.

differences with the soil fractions especially for dry-sieving. Interestingly, bulk soil from dry-sieving grouped closely to bulk soil from wet-sieving and soil fractions. However, the PCoA revealed differences between soil fractions with the wet-sieving method, and high variation between replicates with dry-sieving (Fig. 5). This was confirmed when the ANOSIM was performed without the bulk soil, showing only significant and relatively strong differences between soil fractions when obtained by wet-sieving (R = 0.44, *P* = 0.0001) and no difference with dry-sieving (R = 0.1, *P* = 0.108).

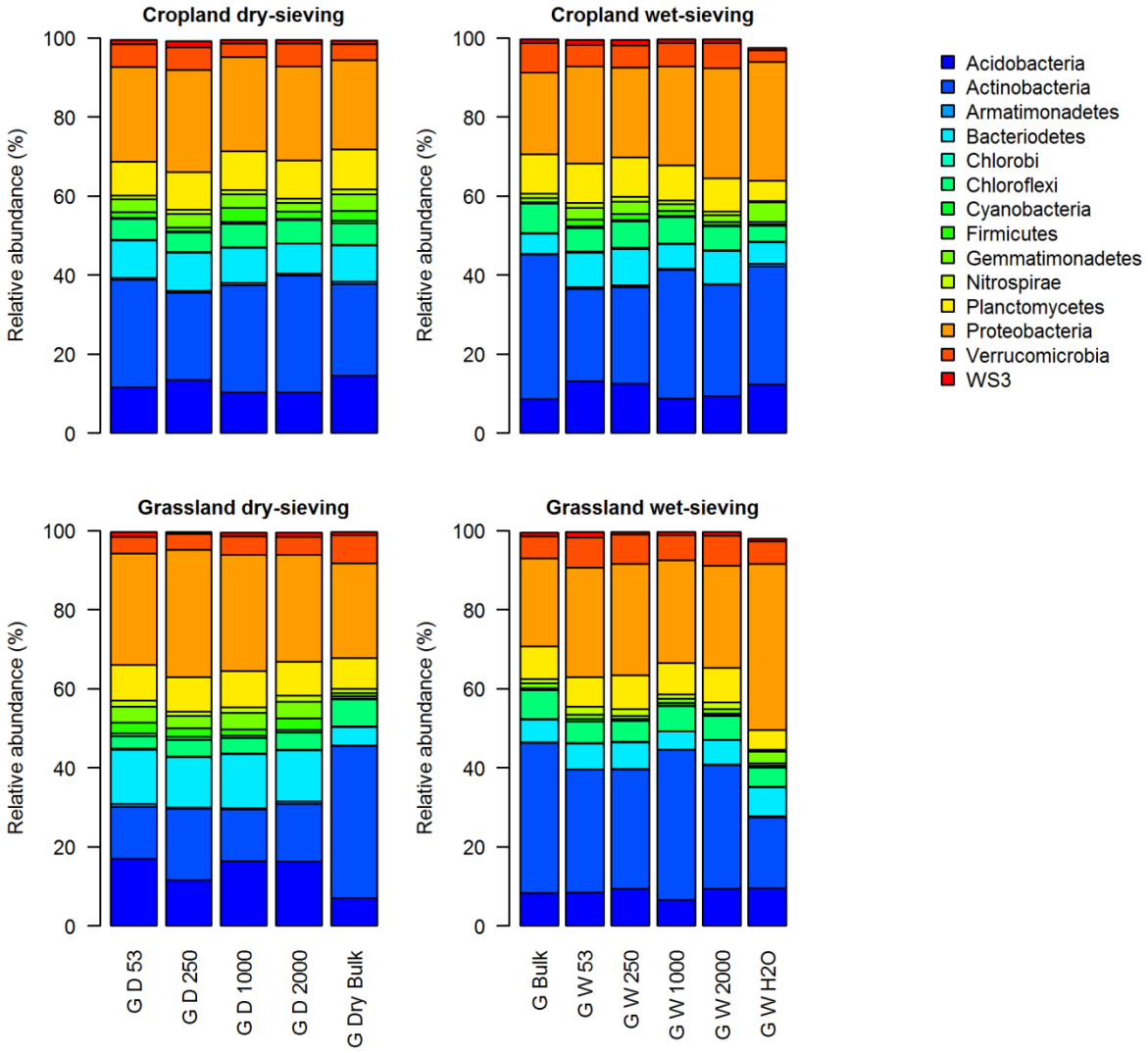

Fig. 6. Relative abundance (%) of bacterial phyla of four soil fractions obtained by dry- or wet-sieving method, bulk soil and water from wet-sieving from cropland and grassland. Means values (*n* = 3) are shown. Only the dominant phyla (~ > 0.2%) are shown.

The relative abundances of most of the dominant phyla were strongly affected by the sieving methods with a decrease with wet-sieving for most of them except for *Actinobacteria, Cyanobacteria* and *Verrucomicrobia* that increased with wet-sieving (Fig. 6; Table S5). The different sizes of soil fractions also affected the relative abundance of most phyla. The differences between sieving methods and soil fractions size were more visible and statistically significant for the grassland than cropland. The differences between cropland and grassland were related to only few of the dominant phyla, with *Chloroflexi*, and *Planctomycetes* that were higher in cropland, while *Nitrospirae*, and *Proteobacteria* were higher in grassland (Fig. 6; Table S5). The water from wet-sieving in grassland showed a significant decrease in *Actinobacteria* and *Planctomycetes* in comparison to the soil fractions, while *Proteobacteria* increased.

PCoA were also performed on the archaeal community composition, showing strong differences between the water from wet-sieving and the rest of the samples although water samples from grassland grouped with the soil fractions (Fig. S2, S3). Then strong differences in archaeal community composition were also found between sieving methods but not between soil fractions.

4. Discussion

The study of the distribution of microbial communities diversity, abundance and activities between different sizes of soil aggregates size classes started more than two decades ago (Chotte et al., 1993; Gupta and Germida, 1988; Jocteur Monrozier et al., 1991; Kanazawa and Filip, 1986; Lensi et al., 1995). The study of microbial distribution in soil aggregates starts from the premise that the vast variation in the size of aggregates, as well as their physico-chemical properties, provides a huge diversity of habitats for microorganisms influencing carbon and nutrients dynamics within the soil. Subsequently, it implies that each soil aggregate size class could harbours specific microbial communities and activities. However, little is known about the effects of size fractionation methods such as sieving on the isolation and interpretation of microbial community data from soil aggregates. The current study clearly shows that dry- or wet-sieving methods affect the acquisition and interpretation of microbial data from different soil aggregates. Furthermore, the effects of sieving methods vary with the site/soil studied, and also which component of the microbial community was studied (i.e. diversity vs. abundance).

Differences in bacterial community composition between sizes of soil aggregates were only revealed in grassland and only when using wet-sieving. Dry-sieving method resulted in high variation between replicates, hindering potential differentiation between sizes. The higher disruption energy introduced by water movement and washing effect provided during wet-sieving in comparison to dry-sieving are likely to be the main factors explaining such differences in the results obtained by both sieving methods (Cambardella and Elliott, 1993; Chotte et al., 1993). This result implies that the different spatial domains of microbial diversity within soil are distinguished by patterns in the adhesive forces within soil that bind organisms, minerals and fluids together. This suggests that some factors that are important in the spatial variation in particle binding to form aggregates may be also important as selective pressures to establish differences in microbial diversity. Similar results were found with the potential enzyme activity, with only the wet-sieving method that revealed significant differences between soil aggregate sizes in comparison to two dry-sieving methods (i.e. soil either air-dry or dry to 10-15% of soil gravimetric water content) (Bach and Hofmockel, 2014). These results highlight the fact that wet-sieving might be a better method over dry-sieving to isolate the different microbial communities within each size fraction, and is relevant for different microbial characteristics: diversity and activity. The bacterial gene abundances showed overall less clear variation between soil aggregates sizes regardless of the sieving methods, although wet-sieving resulted showed more variation in genes abundance between sizes than dry-sieving.

The washing of soil aggregates during wet-sieving did not result in significant cross contamination between aggregate sizes, at least for grassland where significant differences were found. In contrast, dry-sieving and its rubbing effect on the outer part of aggregates may result in stronger cross contamination due to the absence of water carrying the soil particles into the soil fraction < 53 µm, which represent a patchwork of the different soil fractions, and its mass is directly affected by the disruption strength energy (Chotte et al., 1993). This was supported by the high variation between soil fractions replicates for cropland. Relatively low percentages of bacterial genes, often below 1%, were lost in the wet-sieving water, although this percentage was likely to be underestimated. Interestingly, high *narG* gene percentage and relative abundance of *Proteobacteria* in water from grassland may indicate that this gene and phylum might be located in the outer part of the aggregates or inter-aggregates space, where the washing effect was high. In contrast, *Actinobacteria* and *Planctomycetes* decreased in relative abundance, indicating a location within aggregates or high adhesion to soil particles.

Thus, water from wet-sieving might give some indication on the location of some bacterial community.

In cropland no differences in bacterial diversity between aggregate size fractions were found regardless of the sieving method, highlighting that differences between soil aggregates sizes are not always expected but clearly depend on the soil type and land use. Previous studies also showed no difference between microbial community in different size fractions from cropland, likely due to the high turnover of soil aggregates because of anthropogenic activities (e.g. soil tillage, plant harvest) leading to high physical instability of the microenvironment hindering the differentiation of the microhabitats and microbial communities (Blaud et al., 2014). Thus, the absence of difference in microbial diversity between aggregate sizes at a site could potentially be used as an indicator of the instability of the systems and soil health.

Wet-sieving extracted higher gene abundance than dry-sieving. Wetting dry soil was shown to increase the amount of DNA extracted from soil (Clark and Hirsch, 2008), and a physical effect rather than biological might explain the difference for wet-sieving within the ~30 min that the fractionation last. The same trend was found for potential enzyme activity, with four fold greater activity found with wet-sieving in comparison to dry-sieving (Bach and Hofmockel, 2014). This can either reflect an overestimation of the measured variable due to wetting effect (i.e. biological), or accessing a hidden microbial community protected within the pores of the aggregates. In contrast, dry-sieving could lead to under-estimating the microbial gene abundance. Bach and Hofmockel (2014) suggested that wetting the soil leads to over-estimation of potential enzyme activity due to contact between microorganisms and soluble C compounds and the potential short-term microbial metabolic changes. However, there is also a large number of slow growing microorganisms in soil. Most studies showing a rapid response of the microbial community to changes in moisture (< 30 min), were done only on a few microbial strains in optimal laboratory conditions far from *in situ* conditions (Halverson et al., 2000; Lamarre et al., 2008). Nevertheless, the biological effect of wetting on microorganisms cannot be discarded.

Chotte et al. (2002) suggested that studying the microbial community within soil aggregates gives access to changes in microbial community that would not be visible in the bulk soil, and a greater diversity of *Azospirillum*. Most studies assessing microbial community composition within soil aggregates found significant differences with the bulk soil (Blaud et al., 2012; Chotte et al., 2002; Davinic et al., 2012; Ranjard et al., 2000).

Despite the bacterial diversity found not to be higher in each soil aggregates size in comparison to the bulk soil in the study using next generation sequencing (Davinic et al., 2012) and in the current study, taken all together, the different soil aggregates harbour a greater bacterial diversity than the bulk soil. It remains unclear if pooling the same number of DNA extractions from bulk soil as the number of those from soil fractions plus replicates (e.g. 12 DNA extract in the current study) would lead to an increase in bacterial diversity harvested in the bulk soil. This issue could be partly a methodological constraint, as DNA extraction usually uses an extremely small amount of soil; 0.25 g is commonly used, which reduces the representation of the different soil aggregate sizes within the extraction. The recent study from Penton et al. (2016) showed that higher bacterial diversity was found when 10 g of soil was used, which could be related to higher representation of the different soil aggregates sizes and in general the heterogeneous structure of the soil. Similar issue could also be relevant when studying microbial activity that uses often only 1 g of soil (Bach and Hofmockel, 2014)

Overall, it should not be expected that the bulk soil provides a summary of the different soil fractions when working on small amounts of soil. Furthermore, isolated soil fractions are likely to behave in a different way to those *in situ* because of exposure for example to oxygen and high concentration of soil fractions in comparison to their dispersion within a soil horizon. These characteristics could be major limitations when trying to link microbial diversity, abundance and activity between bulk soil and soil fractions, or to model these variables taking into consideration the soil structure.

5. Conclusions

Sieving methods clearly affect the resulting observed bacterial diversity and abundance found in soil aggregates, and there is a need to carefully choose the methods used prior to their study. Wet-sieving was potentially the most adapted method to study microbial community diversity and abundance in soil aggregates in comparison to dry-sieving, although it is time consuming and difficult to perform. Further, studies are needed to assess if wet-sieving is the relevant method across a larger number of land use and soil types, and also to assess if is relevant for the measure of other microbial variables (e.g. RNA). Aggregates isolated with sieving methods are the products of sieving and it might be difficult to relate the microbial results to *in situ* reality. However, aggregates are real units of greater cohesion in the soil formed by biogeochemical processes. Overall, this study raises the question on how

to consider soil structure in the study of soil microbial communities, in order to address important question such as the biological mechanisms controlling soil fertility or the stabilisation or organic matter in soil.

Acknowledgments

This work is supported by the European Commission 7th Framework Programme as a Large Integrating Project, SoilTrEC (www.soiltrec.eu), Grant Agreement No. 244118. We thank Winfried E.H. Blum and Jasmin Schiefer from University of Natural Resources and Life Sciences Vienna (BOKU) for their help with soil sampling in the CZO Marchfeld/Fuchsenbigl.

References

Bach, E.M., Hofmockel, K.S., 2014. Soil aggregate isolation method affects measures of intra-aggregate extracellular enzyme activity. Soil Biology and Biochemistry 69, 54–62. doi:10.1016/j.soilbio.2013.10.033

Beauchamp, E.G., Seech, A.G., 1990. Denitrification with different sizes of soil aggregates obtained from dry-sieving and from sieving with water. Biology and Fertility of Soils 10, 188–193. doi:10.1007/BF00336134

Blaud, A., Chevallier, T., Virto, I., Pablo, A.-L., Chenu, C., Brauman, A., 2014. Bacterial community structure in soil microaggregates and on particulate organic matter fractions located outside or inside soil macroaggregates. Pedobiologia 57, 191–194. doi:10.1016/j.pedobi.2014.03.005

Blaud, A., Lerch, T.Z., Chevallier, T., Nunan, N., Chenu, C., Brauman, A., 2012. Dynamics of bacterial communities in relation to soil aggregate formation during the decomposition of 13C-labelled rice straw. Applied Soil Ecology 53, 1–9. doi:10.1016/j.apsoil.2011.11.005

Cambardella, C.A., Elliott, E.T., 1993. Methods for physical separation and characterization of soil organic matter fractions. Geoderma 56, 449–457. doi:10.1016/0016-7061(93)90126-6

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, 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. doi:10.1038/nmeth.f.303

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 the National Academy of Sciences 108, 4516–4522. doi:10.1073/pnas.1000080107

Chotte, J.-L., Jocteur Monrozier, L., Villemin, G., Albrecht, A., Mulongoy, K., Merckx, R., 1993. Soil microhabitats and the importance of the fractionation method, in: Soil Organic Matter Dynamics and Sustainability of Tropical Agriculture. IITA, Ibadan, pp. 39–45.

Chotte, J.L., Schwartzmann, A., Bally, R., Jocteur Monrozier, L., 2002. Changes in bacterial communities and Azospirillum diversity in soil fractions of a tropical soil under 3 or 19 years of natural fallow. Soil Biology and Biochemistry 34, 1083–1092. doi:10.1016/S0038-0717(02)00041-X

Clark, I.M., Hirsch, P.R., 2008. Survival of bacterial DNA and culturable bacteria in archived soils from the Rothamsted Broadbalk experiment. Soil Biology and Biochemistry 40, 1090–1102. doi:10.1016/j.soilbio.2007.11.021

Clarke, K.R., Green, R.H., 1988. Statistical design and analysis for a “biological effects” study. Marine Ecology-Progress Series 46, 213–226.

Davinic, M., Fultz, L.M., Acosta-Martinez, V., Calderón, F.J., Cox, S.B., Dowd, S.E., Allen, V.G., Zak, J.C., Moore-Kucera, J., 2012. Pyrosequencing and mid-infrared spectroscopy reveal distinct aggregate stratification of soil bacterial communities and organic matter composition. Soil Biology and Biochemistry 46, 63–72. doi:10.1016/j.soilbio.2011.11.012

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10, 996–998. doi:10.1038/nmeth.2604

Fall, S., Nazaret, S., Chotte, J.L., Brauman, A., 2004. Bacterial density and community structure associated with aggregate size fractions of soil-feeding termite mounds. Microbial Ecology 48, 191–199. doi:10.1007/s00248-003-1047-2

Gupta, V.V.S.R., Germida, J.J., 1988. Distribution of microbial biomass and its activity in different soil aggregate size classes as affected by cultivation. Soil Biology and Biochemistry 20, 777–786. doi:10.1016/0038-0717(88)90082-X

Halverson, L.J., Jones, T.M., Firestone, M.K., 2000. Release of Intracellular Solutes by Four Soil Bacteria Exposed to Dilution Stress. Soil Science Society of America Journal 64, 1630. doi:10.2136/sssaj2000.6451630x

Helgason, B.L., Walley, F.L., Germida, J.J., 2010. No-till soil management increases microbial biomass and alters community profiles in soil aggregates. Applied Soil Ecology 46, 390–397. doi:10.1016/j.apsoil.2010.10.002

IUSS Working Group WRB, 2006. World reference base for soil resources 2006, World soil resources rep 103. ed. FAO, Rome.

Jocteur Monrozier, L., Ladd, J.N., Fitzpatrick, R.W., Foster, R.C., Rapauch, M., 1991. Components and microbial biomass content of size fractions in soils of contrasting aggregation. Geoderma 50, 37–62. doi:10.1016/0016-7061(91)90025-O

Kanazawa, S., Filip, Z., 1986. Distribution of microorganisms, total biomass, and enzyme activities in different particles of brown soil. Microbial Ecology 12, 205–215. doi:10.1007/BF02011205

Kandeler, E., Tscherko, D., Bruce, K.D., Stemmer, M., Hobbs, P.J., Bardgett, R.D., Amelung, W., 2000. Structure and function of the soil microbial community in microhabitats of a heavy metal polluted soil. Biology and Fertility of Soils 32, 390–400. doi:10.1007/s003740000268

Kravchenko, A.N., Negassa, W.C., Guber, A.K., Hildebrandt, B., Marsh, T.L., Rivers, M.L., 2014. Intra-aggregate pore structure influences phylogenetic composition of bacterial community in macroaggregates. Soil Science Society of America Journal 78, 1924. doi:10.2136/sssaj2014.07.0308

Lair, G.J., Zehetner, F., Hrachowitz, M., Franz, N., Maringer, F.-J., Gerzabek, M.H., 2009. Dating of soil layers in a young floodplain using iron oxide crystallinity. Quaternary Geochronology 4, 260–266. doi:10.1016/j.quageo.2008.11.003

Lamarre, C., Sokol, S., Debeaupuis, J.-P., Henry, C., Lacroix, C., Glaser, P., Coppée, J.-Y., François, J.-M., Latgé, J.-P., 2008. Transcriptomic analysis of the exit from dormancy of Aspergillus fumigatus conidia. BMC Genomics 9, 417. doi:10.1186/1471-2164-9-417

Lensi, R., Clays-Josser, A., Jocteur Monrozier, L., 1995. Denitrifiers and denitrifying activity in size fractions of a mollisol under permanent pasture and continuous cultivation. Soil Biology and Biochemistry 27, 61–69. doi:10.1016/0038-0717(94)00132-K

Manter, D.K., Vivanco, J.M., 2007. Use of the ITS primers, ITS1F and ITS4, to characterize fungal abundance and diversity in mixed-template samples by qPCR and length heterogeneity analysis. Journal of Microbiological Methods 71, 7–14. doi:10.1016/j.mimet.2007.06.016

McMurdie, P.J., Holmes, S., 2013. phyloseq: An R package for reproducible Interactive analysis and graphics of microbiome census data. PLoS ONE 8, e61217. doi:10.1371/journal.pone.0061217

Mendes, I.C., Bandick, A.K., Dick, R.P., Bottomley, P.J., 1999. Microbial biomass and activities in soil aggregates affected by winter cover crops. Soil Science Society of America Journal 63, 873. doi:10.2136/sssaj1999.634873x

Penton, C.R., Vadakattu, V.S.R.G., Yu, J., Tiedje, J.M., 2016. Size Matters: Assessing Optimum Soil Sample Size for Fungal and Bacterial Community Structure Analyses Using High Throughput Sequencing of rRNA Gene Amplicons. Frontiers in Microbiology 7, 824. doi:10.3389/fmicb.2016.00824

Pylro, V.S., Roesch, L.F.W., Morais, D.K., Clark, I.M., Hirsch, P.R., Tótola, M.R., 2014. Data analysis for 16S microbial profiling from different benchtop sequencing platforms. Journal of Microbiological Methods 107, 30–37. doi:10.1016/j.mimet.2014.08.018

R Development Core Team, 2015. R: a language and environment for statistical computing.

Ranjard, L., Nazaret, S., Gourbiere, F., Thioulouse, J., Linet, P., Richaume, A., 2000. A soil microscale study to reveal the heterogeneity of Hg(II) impact on indigenous bacteria by quantification of adapted phenotypes and analysis of community DNA fingerprints. FEMS Microbiology Ecology 31, 107–115. doi:10.1111/j.1574-6941.2000.tb00676.x

Ririe, K.M., Rasmussen, R.P., Wittwer, C.T., 1997. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Analytical Biochemistry 245, 154–160. doi:10.1006/abio.1996.9916

Sainju, U.M., 2006. Carbon and nitrogen pools in aoil aggregates seperated by dry and wet sieving methods. Soil Science 171, 937–949. doi:10.1097/01.ss0000228062.30958.5a

Sainju, U.M., Terrill, T.H., Gelaye, S., Singh, B.P., 2003. Soil aggregation and carbon and nitrogen pools under Rhizoma peanut and perennial weeds. Soil Science Society of America Journal 67, 146–155.

Schutter, M.E., Dick, R.P., 2002. Microbial community profiles and activities among aggregates of winter fallow and cover-cropped soil. Soil Science Society of America Journal 66, 142–153.

Seech A.G. and Beauchamp E.G., Denitrification in soil aggregates of different sizes, Soil Sci. Soc. Am. J. 52, 1988, 1616–1621.

Sessitsch, A., Weilharter, A., Gerzabek, M.H., Kirchmann, H., Kandeler, E., 2001. Microbial population structures in soil particle size fractions of a long-term fertilizer field experiment. Applied and Environmental Microbiology 67, 4215–4224. doi:10.1128/AEM.67.9.4215-4224.2001

Sey, B.K., Manceur, A.M., Whalen, J.K., Gregorich, E.G., Rochette, P., 2008. Small-scale heterogeneity in carbon dioxide, nitrous oxide and methane production from aggregates of a cultivated sandy-loam soil. Soil Biology and Biochemistry 40, 2468–2473. doi:10.1016/j.soilbio.2008.05.012

Tsiknia, M., Tzanakakis, V.A., Paranychianakis, N.V., 2013. Insights on the role of vegetation on nitrogen cycling in effluent irrigated lands. Applied Soil Ecology 64, 104–111. doi:10.1016/j.apsoil.2012.10.010

Vaisanen, R.K., Roberts, M.S., Garl, J.L., Frey, S.D., Dawson, L.A., 2005. Physiological and molecular characterisation of microbial communities associated with different water-stable aggregate size classes. Soil Biology and Biochemistry 37, 2007–2016. doi:10.1016/j.soilbio.2005.02.037

Yoder, R.E., 1936. A direct method of aggregate analysis of soils and a study of the physical nature of erosion losses. Journal of American Society of Agronomy 28, 337–351.

Supplementary information

Table S1. Description of the primers used to target each community and the annealing temperature of each Q-PCR assays.

Target gene	Primer	Sequence 5'-3'	Annealing temp. (°C) and time (s)	References
Bacterial	519F	GCCAGCAGCCGCGGTAAT	58 (30 s)	Lane, 1991
16SrRNA	907R	CCGTCAATTCCTTTGAGTTT		Stubner and Meuser, 2000
Archaeal	Arch 0025F	CTGGTTGATCCTGCCAG	58 (30 s)	Vetriani et al., 1999
16SrRNA	Arch 364R	ACGGGGCGCACGAGGCGCGA		Vetriani et al., 1999
Fungal	ITS1f	TCCGTAGGTGAACCTGCGG	50 (45 s)	Gardes and Bruns, 1993
ITS	5.8s	CGCTGCGTTCTTCATCG		Vilgalys and Hester, 1990
nifH	nifHF	AAAGGYGGWATCGGYAARTCCACCAC	62.5 (60 s)	Rösch and Bothe, 2005
	nifHRb	TGSGCYTTGTCYTCRCGGATBGGCAT		Rösch and Bothe, 2005
amoA	amoA_F	GGHGACTGGGAYTTCTGG	55.3 (30 s)	Holmes et al., 1995
Bacteria	amoA_R	CCTCKGSAAAGCCTTCTTC		Okano et al., 2004
amoA	amoAF	STAATGGTCTGGCTTAGACG	55 (35 s)	Francis et al., 2005
Archaea	amoAR	GCGGCCATCCATCTGTATGT		Francis et al., 2005
narG	NARG F	TCGCCSATYCCGGCSATGTC	63 (30 s)	López-Gutiérrez et al., 2004
	NARG R	GAGTTGTACCAGTCRGCSGAYTCSG		López-Gutiérrez et al., 2004
nirS	NIRS4Q F	GTSAACGYSAAGGARACSGG	63 (30 s)	Braker et al., 1998
	NIRS6Q R	GASTTCGGRTGSGTCTTSAYGAA		Braker et al., 1998
nosZ	nosZ1840_F	CGCRACGGCAASAAGGTSMSSGT	67 (30 s)	Henry et al., 2006
	nosZ2090_R	CAKRTGCAKSGCRTGGCAGAA		Henry et al., 2006

Table S2. Overview table of the ANOVA of the aggregate distribution with sites, soil fractions and sieving methods as factors. Significant P values ($P < 0.05$) are shown in bold.

Factors	F values	*P* value
Sites	0.04	0.84
Fractions	**31.93**	**9.64 10-^{10}**
Sieving	0.07	0.79
Sites:fractions	**5.45**	**0.004**
Sites:sieving	0.0043	0.95
Fractions:sieving	**26.12**	**9.83 10^{-9}**
Sites:fractions:sieving	**8.65**	**0.00024**

Table S3. Overview table of the ANOVA of the relative abundance of microbial genes, with sites, soil fractions and sieving methods as factors. Significant P values ($P < 0.05$) are shown in bold.

Factors		Bacteria	Fungi	*nifH*	AOB	AOA	*narG*	*nirS*	*nosZ*
Sites	F value	**296.87**	**65.35**	**277.08**	**116.70**	**6.36**	0.69	**191.4**	**147.83**
	P values	**<2 10-16**	**6.09 10-10**	**<2 10-16**	**1.99 10-13**	**0.016**	0.41	**<2 10-16**	**1.14 10-14**
Fractions	F value	**3.06**	0.55	1.01	0.94	0.88	0.49	1.06	0.73
	P values	**0.027**	0.70	0.41	0.45	0.49	0.75	0.39	0.57
Sieving	F value	0.996	1.35	3.67	**12.66**	1.36	3.34	**18.28**	**10.07**
	P values	0.324	0.25	0.06	**0.00098**	0.25	0.07	**0.0001**	**0.003**
Sites:fractions	F value	2.49	1.24	0.37	1.29	0.86	2.49	0.43	0.86
	P values	0.059	0.31	0.82	0.29	0.50	0.06	0.78	0.50
Sites:sieving	F value	0.52	1.90	0.38	0.03	0.24	0.07	0.0007	0.50
	P values	0.47	0.17	0.54	0.86	0.62	0.79	0.98	0.48
Fractions:sieving	F value	1.45	0.53	0.16	0.19	0.60	0.062	0.18	0.079
	P values	0.24	0.71	0.96	0.94	0.67	0.99	0.94	0.99
Sites:fractions:sieving	F value	0.41	1.26	1.37	0.87	1.60	0.58	0.56	0.59
	P values	0.80	0.30	0.26	0.49	0.19	0.68	0.69	0.67

Table S4. Overview table of the ANOVA of the microbial gene express as percentage of the bacterial 16S rRNA gene copies, with sites, soil fractions and sieving methods as factors. Significant P values ($P < 0.05$) are shown in bold.

Factors		*nifH/16S rRNA*	AOB/ 16S rRNA	*narG/16S rRNA*	*nirS/16S rRNA*	*nosZ/16S rRNA*
Sites	F value	**33.47**	**1391.1**	**34.31**	2.56	0.01
	P values	**9.50 10-7**	**<2.2 10-16**	**7.49 10-7**	0.12	0.91
Fractions	F value	**8.13**	**6.92**	0.85	**11.56**	**8.57**
	P values	**6.80 10-5**	**0.00025**	0.5	**2.48 10-6**	**4.94 10-5**
Sieving	F value	**4.44**	**19.63**	0.14	**47.6**	**26.84**
	P values	**0.041**	**7.13 10-5**	0.71	**2.58 10-8**	**7.54 10-6**
Sites:fractions	F value	2.33	4.86	1.61	2.59	**8.36**
	P values	0.07	0.0027	0.19	0.051	**6.12 10-5**
Sites:sieving	F value	**4.95**	0.32	2.8	1.026	0.27
	P values	**0.03**	0.58	0.1	0.32	0.61
Fractions:sieving	F value	**3.06**	3.2	0.44	**5.17**	**3.88**
	P values	**0.027**	0.02	0.78	**0.0019**	**0.01**
Sites:fractions:sieving	F value	**3.71**	1.4	0.15	2.51	**3.64**
	P values	**0.01**	0.24	0.96	0.57	**0.01**

Table S5. Overview table of the ANOVA of the relative abundance of bacterial phylum, with sites, soil fractions and sieving methods as factors. Significant P values ($P < 0.05$) are shown: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Factors	Sites	Fractions	Sieving	*Sites:fractions*	*Sites:sieving*	*Fractions:sieving*	*Sites:fractions:sieving*
Acidobacteria		*	***	**	*		***
Actinobacteria		***	***	*	**		**
Armatimonadetes			***		*		***
Bacteriodetes		*	***				*
Chlorobi		*					
Chloroflexi	*	***	***	*			
Cyanobacteria			***		**		*
Firmicutes		*	***				***
Gemmatimonadetes			***				*
Nitrospirae	***	***		**			
Planctomycetes	***	**					
Proteobacteria	***	***					
Verrucomicrobia			***				**
WS3				*			

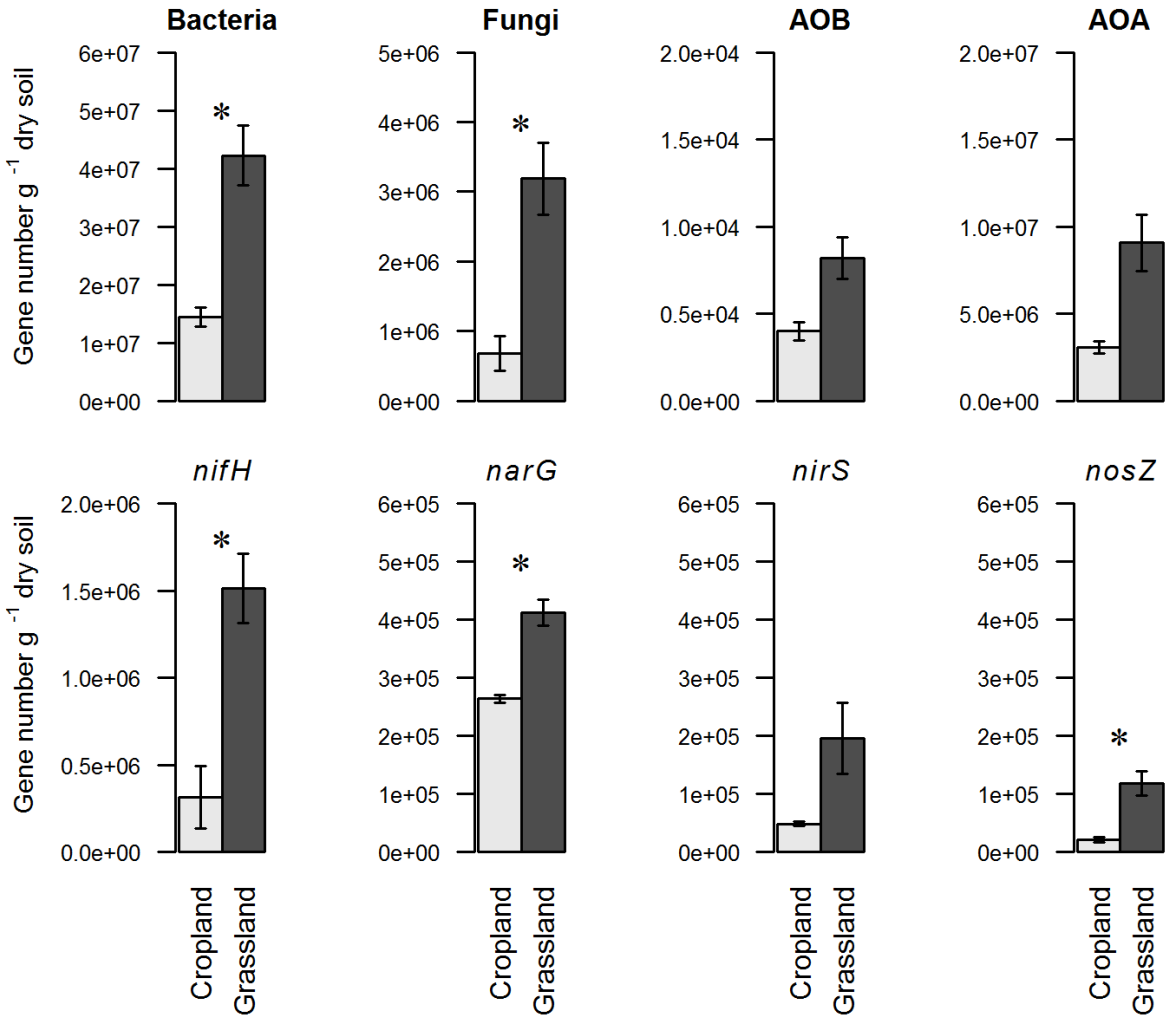

Fig. S1. Variation in microbial gene abundance of water from wet sieving method. All abundances are expressed on the basis of 1 g of dry mass of soil fractionated. Means values ± standard error (n = 3) are shown. * indicate significant ($P < 0.05$) difference between cropland and grassland for a specific gene. AOB: amoA bacteria. AOA: amoA archaea.

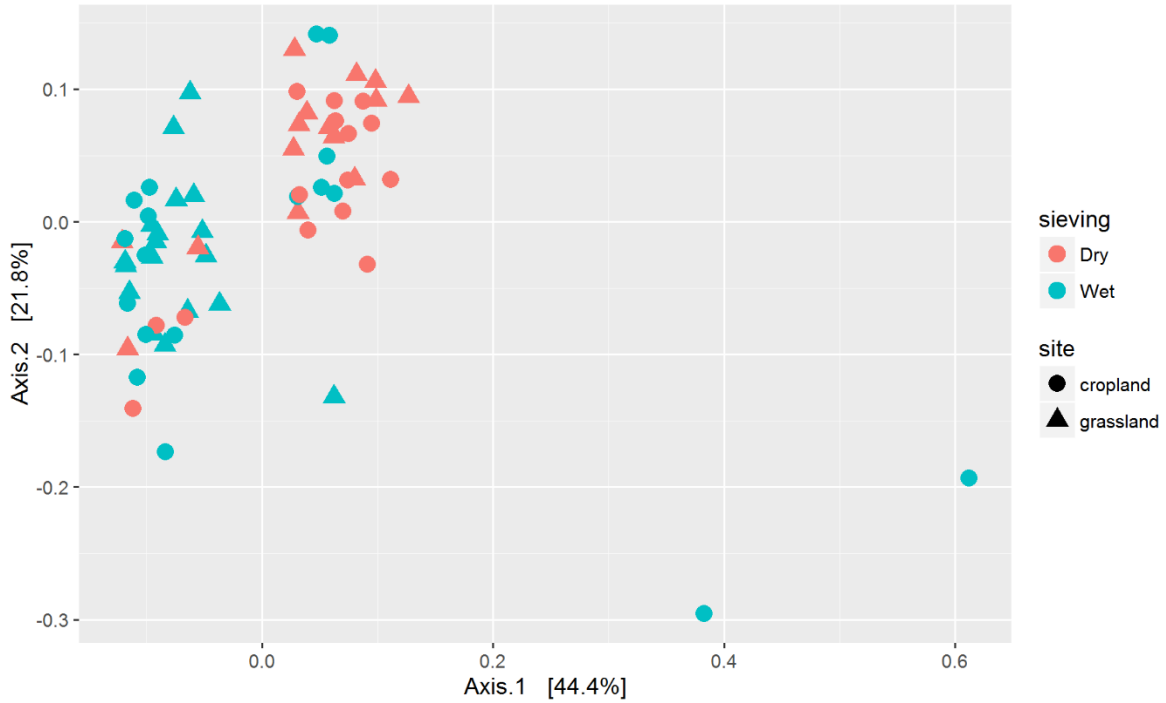

Fig. S2. PCoA of archaeal community of four soil fractions obtained by dry- or wet-sieving method and bulk soil from cropland and grassland. The PCoA was based on relative abundance of OTU and generated using Bray-Curtis distance. The samples isolated from the rest of the samples correspond to water from the wet-sieving.

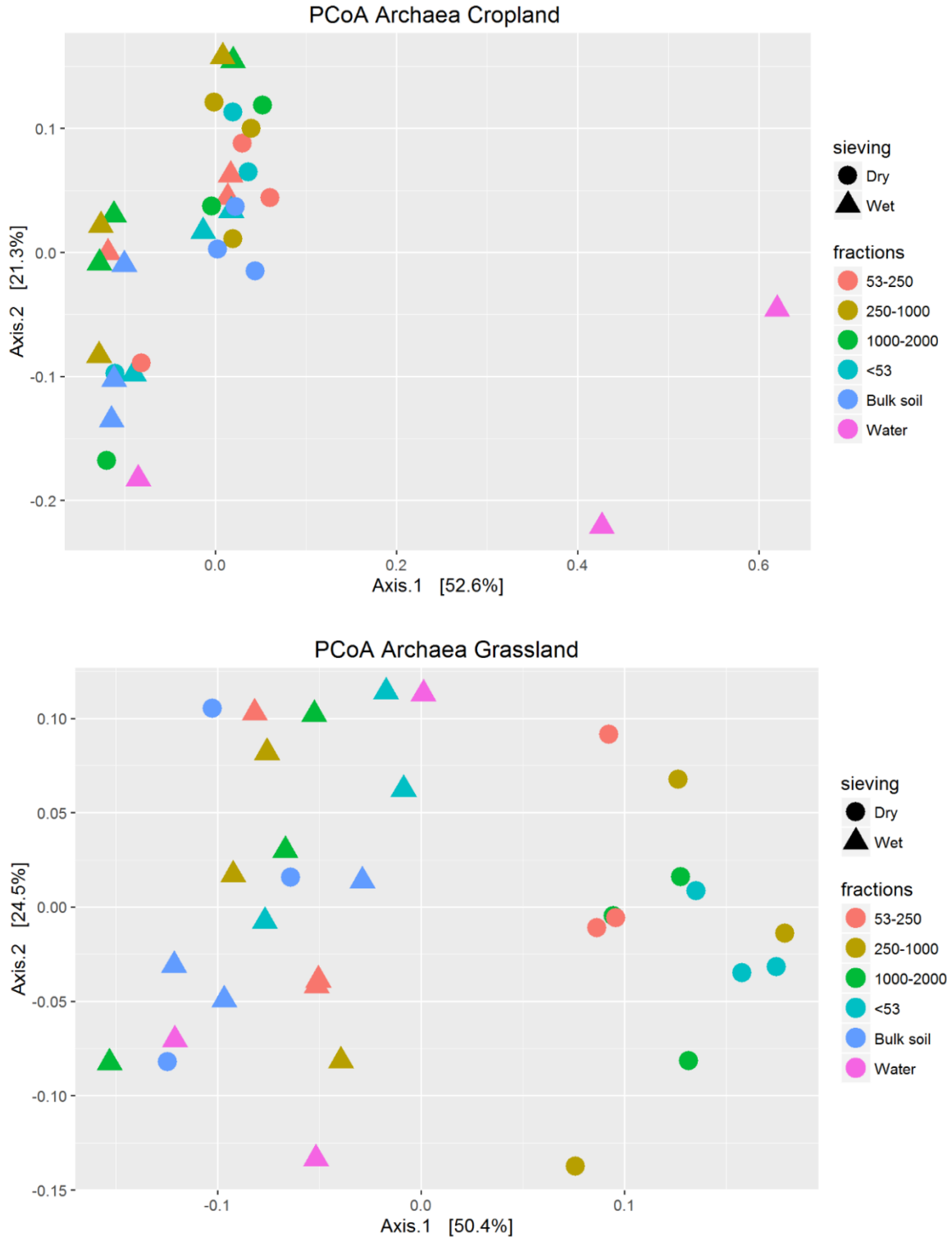

Fig. S3. PCoA of archaeal community of four soil fractions obtained by dry- or wet-sieving method and bulk soil from cropland (top) and grassland (bottom). The PCoA were based on relative abundance of OTU and generated using Bray-Curtis distance.