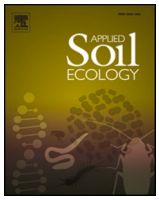
Applied Soil Ecology 200 (2024) 105464

Contents lists available at ScienceDirect

Applied Soil Ecology

journal homepage: www.elsevier.com/locate/apsoil



Research paper

Profiling the eukaryotic soil microbiome with differential primers and an antifungal peptide nucleic acid probe (PNA): Implications for diversity assessment

Haotian Wang [a], Kenneth Dumack [b], Daniel V. Rissi [a], Damien R. Finn [a], Michael Bonkowski [b], Christoph C. Tebbe [a,*]

[a] Thünen Institute of Biodiversity, Braunschweig, Germany
[b] Terrestrial Ecology, Institute of Zoology, University of Cologne, Köln, Germany

ARTICLE INFO

Keywords:
Soil DNA
Peptide nucleic acid (PNA)
Fungal diversity
Protist diversity
18S rRNA genes
Universal eukaryotic PCR primers

ABSTRACT

The analysis of 18S rRNA gene amplicons is an important tool to characterize the diversity of the eukaryotic soil microbiome. Here we analyzed two primer sets (TAReuk, EKeuk) and the impact of a newly designed antifungal peptide nucleic acid to enhance the detection of protists *in silico* and with soil DNA extracted from croplands and a forest. The *in silico* analyses showed for TAReuk pronounced specificities for protist SAR supergroup and Metazoa, while EKeuk was particularly specific for Ascomycota and Basidiomycota. *In silico*, the PNA matched with the majority of Ascomycota (81.3 %) and Basidiomycota (65.4 %), but with <6 % of protists. The contrasting primer specificities were confirmed with soil DNA, but the proportion of protist amplicons was similar. In contrast to *in silico*, effects of the PNA were not as clear with soil DNA, even though it completely inhibited the amplification of the targeted fungal sequences. PNA effects were more pronounced with TAReuk, and results with cropland and forest soil DNA were not consistent, *e.g.*, for cropland, PNA decreased the relative abundance of fungi but for forest it was the opposite, possibly because of different fungal diversity. The divergence between PNA *in silico*-predictions and results with soil DNA are likely an outcome of primer binding to <100 % complementary target sequences and a still limited DNA sequence databases for soil microbial eukaryotes. With TAReuk, the presence of PNA enhanced the detection of Conosa and, thus, could be a useful tool to study this group in the future.

1. Introduction

Cultivation-independent analyses of PCR-amplified 18S rRNA genes sequences from environmental DNA provides efficient and sensitive means for characterizing the composition of complex eukaryotic microbiomes, *e.g.*, in marine, fresh water, or terrestrial ecosystems (Almela et al., 2023; Catlett et al., 2020; Simon et al., 2015a). PCR primers binding to phylogenetically conserved chromosomal gene regions flanking hypervariable regions of the 18S rRNA gene, so-called universal primers, allow capturing a high diversity of eukaryotes. Bioinformatic analyses of the sequences of the hypervariable regions can then be used to search for sequence similarity in databases and thereby provide information on their taxonomic identity (Hugerth et al., 2014). Just like 16S rRNA genes of prokaryotes (and the mitochondria and chloroplasts of eukaryotes), 18S rRNA genes carry a total of 9 phylogenetically hypervariable regions, designated V1 to V9, which are separated by the potential binding sites of universal primers (Banos et al., 2018; Xie et al., 2011). The total size of the 18S rRNA gene for most eukaryotic phyla is in the range of 1800 to 1900 bp, but the extremes are 1500 to about 4500 bp (Holzer et al., 2007; Xie et al., 2011). This length exceeds the size of PCR amplicons used for the currently most popular DNA sequencing approach with Illumina MiSeq technology, which is in the range of 300 to 600 bp. Thus, for studying eukaryotic communities, choices must be made for the selection of both universal primers and the targeted hypervariable regions (Vaulot et al., 2022).

However, variation of 18S rRNA genes reveals that Eukaryote phyla differ significantly in their “conserved regions”, just as different hypervariable regions are not equally variable within (comparing V regions to each other) and between organisms (differences of the same V region

* Corresponding author at: Thünen Institut für Biodiversität, Bundesallee 65, 38116 Braunschweig, Germany.
E-mail address: christoph.tebbe@thuenen.de (C.C. Tebbe).

https://doi.org/10.1016/j.apsoil.2024.105464
Received 31 January 2024; Received in revised form 21 May 2024; Accepted 29 May 2024
Available online 2 June 2024

between taxa), respectively. Therefore, the selection of both primers and hypervariable regions has an impact on the actual diversity of amplicon sequence variants (ASV) retrieved from an environmental DNA (Taib et al., 2013; Zhao et al., 2019). To obtain a more robust picture of the eukaryotic community composition, single studies have therefore included an assessment of more than one 18S rRNA gene region. For example, a eukaryotic community in marine anoxic water was assessed by targeting in parallel both the V4 and the V9 regions (Stoeck et al., 2010), and microeukaryotes of a river were characterized by merging amplicon libraries of V1-V2 and V4 regions (Choi et al., 2022).

For soil DNA, 18S rRNA gene amplicon libraries typically contain a mixture of sequences from fungi, protists, plants and animals (Du et al., 2022; Lentendu et al., 2014; Zhao et al., 2019). Because of their greater biomass, typically fungal amplicons predominate in such libraries (Kramer et al., 2016). The relative abundance of sequences from these groups is not a true indicator of absolute abundance *in situ*, because primers will not amplify these groups with equal efficiency. For the specific detection of fungi from soil DNA, primers targeting the ITS1 or ITS2 regions are therefore widely used, thus avoiding co-amplification of other eukaryotic sequences (Yu et al., 2022). However, ITS are more variable than 18S rRNA genes and, thus, increase the risk that primers may fail to amplify some fungal groups (Yarza et al., 2017). For protists, on the other hand, defined as unicellular eukaryotes which are neither fungi, nor plants nor animals, there is no alternative universal primer pair, since this group is polyphyletic (Adl et al., 2019; Pawlowski et al., 2012). Targeting the 18S rRNA gene should thus be, despite the disadvantage of PCR-co-amplifying fungi, plants and metazoan, a way forward to characterize a broad diversity of soil protists (Santos et al., 2020; Singer et al., 2021). However, if designed to be "universal", 18S rRNA gene primers may miss parts or entire phylogenetic linages, especially Amoebozoa and Heterolobosea (Geisen et al., 2014), because of their comparably larger amplicon sizes which tend to amplify less with PCR than shorter sequences. This bias can be partially circumvented by targeting different phylogenetic linages with specific primers (Vaulot et al., 2022). Furthermore, the detection of less abundant taxa may be masked by preferential amplification of dominant groups (Pascoal et al., 2021). However, results obtained with separate PCR primers make it more difficult to directly compare the co-occurrence of different protist groups.

The use of both blocking primers and peptide nucleic acid (PNA) oligonucleotides facilitate an effective and targeted suppression of specific undesired DNA amplicon sequences when using universal primers (Lefevre et al., 2020; von Wintzingerode et al., 2000). Blocking primers, which do not allow elongation during PCR because of C3 spacers at the 3′ end, have been designed to inhibit the co-amplification of sequences of various protist groups (Alveolata, Rhizaria, Stramenopiles) for investigating fungal community compositions (Banos et al., 2018). Other blocking primers were used to inhibit the amplification of metazoan sequences when studying protist community compositions in ocean surface water (Tan and Liu, 2018). PNA probes, which form more stable associations to complementary DNA sequences than corresponding DNA oligonucleotides (Nielsen, 2004), were applied to inhibit the amplification of 16S rRNA gene sequences of mitochondria and chloroplast when analyzing plant or animal associated bacterial communities (Bell et al., 2021; Fitzpatrick et al., 2018; Jackrel et al., 2017). A PNA probe was also used to suppress amplification of plant DNA when studying protistan communities in the rhizosphere of maize (Taerum et al., 2020). Due to these studies, we suspected that the use of a PNA probe could be an option to enhance the detection of protist diversity by suppressing the co-amplification of fungal sequences from soil DNA.

The objective of this study was to develop an antifungal PNA probe and determine its efficiency for enhancing the detection of protists from soil DNA. We evaluated its performance with two primer pairs ("TAReuk" and "EKeuk") both targeting the V4 region of the 18S rRNA gene. The primer pair TAReuk had been developed to study protist communities in marine water samples where fungal sequences pose no problems (Stoeck et al., 2010). However, these primers have since been applied in studies on protist communities from different environmental substrates, including soil (Almela et al., 2023; Du et al., 2022; Zhao et al., 2019). Primer pair EKeuk was developed to characterize protist communities from shallow freshwater systems (Simon et al., 2015a) with a reverse primer specifically biased against sequences of Metazoa (Bower et al., 2004).

Here, we developed the antifungal PNA probe based on *in silico* analyses and tested both primer pairs in presence and absence of the antifungal PNA probe with soil DNA originating from two case studies: (1) A comparison of cropland and forest soils of immediate vicinity (variation in land use), and (2) cropland soils from three neighboring fields with variation in soil texture (loam *vs.* silt loam) and management (conserved *vs.* conventional tillage). We hypothesized that (i) both primers would reveal different relative abundances of fungal and protist taxa, with EKeuk less efficiently amplifying 18S rRNA gene sequences from Metazoa and (ii) that the presence of the antifungal PNA probe would strongly reduce the relative abundance of fungal sequences in the emerging amplicon libraries, irrespective of whether TAReuk or EKeuk were used. Our overall intention was to evaluate the efficacy of using alternative primers and an antifungal PNA to broaden the detection of eukaryotic microorganisms, and especially the diverse phylogenetic linages of protists from soil DNA.

2. Materials and methods

2.1. Design of the peptide nucleic acid (PNA) probe

The anti-fungal PNA probe was designed to target sequences from Dikarya (Ascomycota and Basidiomycota), the most common soil fungi. Firstly, the sequences belonging to Dikarya were obtained from the SILVA 138 database (Quast et al., 2013) and aligned using Clustal Omega (Sievers and Higgins, 2014). The tool DEGEPRIMER was used to optimize the alignment in order to identify PNA candidates (Hugerth et al., 2014). PNA probe candidates were then tested using Mothur (pcr.seqs function) against the V4 region of the 18S rRNA gene of Dikarya and protists (Schloss et al., 2009). The *in silico* selection of PNA candidates was based on the ratio: (number of sequences matching Dikarya / number of sequences matching the target protists) / total of V4 reads. Selected PNA candidates were further tested using the tool TestProbe from the SILVA database to validate coverage of the antifungal PNA probe against the complete 18S rRNA database (Quast et al., 2013). The final probe, selected for this study, was 12 bp in length with a 5′-3′ sequence TCRGCACCTTAC. The PNA probe was custom-synthesized at 90 % purity (Panagene, Daejeon, Republic of Korea). A total of 50 nmol was dissolved upon arrival in 1 mL sterile double distilled water to make a final stock solution of 50 μM.

2.2. Properties of soils

The soils of this study originated from two case studies in Lower Saxony, Northern Germany, near Hildesheim. Case study 1 was a comparison of land use, *i.e.* cropland *vs.* forest soils taken in close vicinity from each other, Case study 2 was a comparison of three cropland soils in immediate vicinity to each other, but with different texture and tillage. The Case study 1 site is a chernozem silt loam located near Harsum (52°12′06″ N; 9°59′52″ E; 86 m a.s.l.) and the Case study 2 site is a luvisol soil near Adenstedt (all *ca.* 80 m a.s.l.) with three treatments, including one clay soil with conservative tillage (abbreviated **CS**, 52°00′29.2″ N; 9°55′41.9″ E), one loam with conservation tillage (**LS**, 52°00′27.4″ N; 9°56′13.6″ E) and one loam with conventional tillage (**LV**, 52°00′29.2″ N; 9°56′22.6″ E). The relative proportions of clay, silt and sand were 31.9 %, 63.3 % and 4.8 % (CS), 17.1 %, 80.7 % and 2.2 % (LS), and 17.2 %, 80.8 % and 2.0 % (LV), respectively. The cropland soils in Case study 2 were collected from the upper 10 cm of the soil profile and replicated in time. Samples were collected on April 14, July 21, and

October 13, all in 2020, with three spatial replicates for each date. Soils from Case study 1 were collected in March 2019 (kindly provided by Julia Schroeder, Thünen Institute for Climate Smart Agriculture). Properties of these soils are described elsewhere (Schroeder et al., 2021). Upon arrival in the laboratory, all soil samples were adjusted with water to *ca.* 60 % water holding capacity, (2 mm mesh size) and stored at −80 °C until use.

2.3. Soil DNA extraction, PCR, and DNA sequencing

Soil DNA was extracted from 0.5 g samples using the FastDNA®SPIN Kit (MP Biomedicals, Eschwege, Germany) with lysing matrix E and using their bead-beating device (MPFastPrep Instrument) set at 6 m s^{-1} for 40 s, following the manufacturer's procedures. DNA concentrations were determined by the Quant-iT™ Picogreen dsDNA assay kit (Thermo Fisher Scientific, Osterode, Germany) using a fluorescence spectral photometer (Mithras LB 940; Berthold Technologies (Bad Wildbad, Germany). All soil DNA solutions were adjusted to 10 ng μL^{-1}.

The 18S rRNA gene target sequences were amplified with a two-step PCR protocol (Finn et al., 2022). All reactions were conducted in a Mastercycler X50s (Eppendorf, Hamburg, Germany). For the first step, 1 μL of template DNA was added to a final volume of 25 μL PCR mix composed of 5 × Q5 reaction buffer and final concentration of 0.2 mM of dNTPs, 0.5 μM of each primer, and 0.02 U μL^{-1} of the Q5 High-Fidelity DNA polymerase (0.02 U Ll^{-1}) (New England Biolabs, Ipswich, MA, USA). For PCR with the PNA probe, reaction mixtures included the same concentrations as the primers (0.5 μM). Two primer sets, both amplifying the V4 hypervariable region of the 18S rRNA gene were selected utilizing the PR^2 and the pr2-primers databases (Guillou et al., 2013; Vaulot et al., 2022). The first primer pairs were TAReuk454FWD1 (5′-CCAGCASCYGCGGTAATTCC-3′) and TAReukREV3 (5′-ACTTTCGTTCTTGATYRA-3′) (Stoeck et al., 2010). This first step of PCR was conducted under the following thermocycling conditions: an initial denaturation step at 95 °C for 2 min, then 24 cycles of 95 °C for 30 s, 52 °C for 30 s, and 72 °C for 30 s; followed by a final extension at 72 °C for 5 min. The second primers were EK-565F (5′-GCAGTTAAAAAGCTCGTAGT-3′) (Simon et al., 2015a) and 18S-EUK-1134-R (5′-TTTAAGTTTCAGCCTTGCG-3′) (Bower et al., 2004; Simon et al., 2015a). This first step of PCR was conducted under the following thermocycling conditions: an initial denaturation step at 95 °C for 2 min, then 24 cycles of 95 °C for 30 s, 58 °C for 45 s, and 72 °C for 90 s; followed by a final extension at 72 °C for 10 min.

A total of 1 μL of the resulting PCR solution was then diluted 10-fold in nuclease-free water (New England Biolabs) and from the resulting template solution, two parallel second PCRs were started with 1 Ll PCR product each. PCR reagents and cycling conditions were identical to the first PCR, except that the 18S rRNA gene primers were replaced by barcoded primers for Illumina sequencing (IDT, Integrated DNA Technologies, Coralville, IA, USA). No PNA treatments were applied for this PCR. The parallel PCRs were combined and agarose electrophoresis (2.0 % wt/vol in TBE) was used to confirm the size and quality of the PCR products (Sambrook, 2001).

A total of 25 μl of PCR amplicons from each sample were then purified and normalized using SequalPrep Normalization Plate Kit (Invitrogen GmbH, Karlsruhe, Germany) to reach a concentration of 1–2 ng μL^{-1} for each sample. All purified soil DNA was pooled to construct a sequencing library. Ovation Rapid DR Multiplex System 1–96 (NuGEN Technologies, Redwood City, CA, USA) was used to prepare libraries, and standard Illumina adapters and sequencing primers were ligated using MyTaq (Bioline, Luckenwalde, Germany). Paired-end (2 × 300 bp) sequencing was performed on the Illumina MiSeq platform by LGC Genomics (Berlin, Germany).

2.4. Bioinformatic analysis

The QIIME2 platform (Bolyen et al., 2019) was used to process the DNA sequencing data. Sequences were reoriented into forward or reverse direction where necessary by an in-house Python script (github.com/DamienFinn/MiSeq_read_reorientation). Cutadapt paired-end methods were performed to demultiplex the reoriented reads (Martin, 2011). Afterwards, vsearch join-pairs function was used to merge the forward and reverse reads following q-score-joined method (Rognes et al., 2016). After the merge step, sequence reads were truncated at positions 600 and 40 by DADA2 denoise-single function (Callahan et al., 2016). ASV were considered as biologically meaningful, unique PCR-amplified sequences capable of differing by as little as a single nucleotide (Callahan et al., 2017). The taxonomy of the 18S rRNA gene-defined ASV was assigned using the Protist Ribosomal Reference (PR2) database (version_4.14.0) https://pr2-database.org/ (Laure et al., 2013), with a minimum bootstrap confidence threshold of 80 % (Guillou et al., 2012). For TAReuk (TAReuk454FWD1/TAReukREV3), a total of 10,854,988 raw reads and for EKeuk (EK-565F/EUK-1134R), a total of 8,384,010 raw reads were obtained from the 66 samples included in this study. These were further analyzed by the DADA2 pipeline. These analyses were done separately for each of the two case studies (land use and cropland). All quality-filtered DNA sequences amplified in this study were deposited at the European Nucleotide Archive (PRJEB70390).

2.5. Statistical analyses

Analysis of variance (ANOVA) was firstly conducted to test any significant differences between the absence and presence of PNA probe groups using the SPSS Statistics 20 (SPSS Inc., Chicago, IL, USA). To ensure the validity of the statistical test results, the assumptions of normality and homogeneity of variances (heteroscedasticity) were checked using the Shapiro-Wilk test and Levene's test, respectively. Then, if the assumptions of normality and homogeneity of variances (heteroscedasticity) were met, statistical significances for diversity indices between the absence and presence of PNA probe were investigated by Duncan's multiple comparison test. Good's coverage scores were calculated in R (version 4.1.2) based on the criteria of Good's coverage = 1 − (sum of singleton ASV / sum of all ASV) (Yuan et al., 2018). Alpha-diversity, *i.e.* Shannon index (H) and richness (S) of fungal and protistan communities were derived in R (version 4.1.2) using the Vegan package (Yuan et al., 2018). Principal component analysis (PCA) was performed from centre-log ratio transformed relative abundance of protistan and fungal communities with the “provenance” package in R (Gloor et al., 2016). We used permutational multivariate analysis of variance (PERMANOVA) to test the Bray-Curtis distances among the differences of community structures for soil protistan and fungal dissimilarity matrices. This included the analysis of differences caused by presence of PNA probe and also by different cropland sites with the adonis function in the vegan R package (ver. 2.6–4) (Oksanen, 2010).

3. Results

3.1. In silico specificity analyses of the antifungal PNA probe and the selected primer sets

The TestProbe function of the SILVA database indicated that 63.5 % of all fungal sequences would perfectly match the newly designed PNA probe (Table 1). Specifically, it matched with 81.3 % of sequences of Ascomycota and 65.4 % of Basidiomycota, respectively. The match for other fungal phyla was much lower. For protists, the PNA probe matched perfectly only to 0.1 % of the sequences of Cercozoa, a group suspected to be abundant in cropland soils (Oliverio et al., 2020). For all other groups, including the SAR (Stramenopiles, Alveolata, Rhizaria) supergroup and Amoebozoa, matches were below 6 %. Furthermore, the PNA probe only matched to 0.5 % of the 18S rRNA gene sequences of Metazoa and 0.1 % Streptophyta. Thus, overall, the *in silico* analyses indicated a high potential of our newly designed PNA probe for reducing of PCR amplified Ascomycota and Basidiomycota 18S rRNA gene sequences

Table 1
In silico evaluation of the specificity of the PNA probe of this study, and two primer pairs TAReuk and EKeuk, as evaluated by the SILVA test probe and prime functions.

Classification Taxonomy		Total sequences	Matched sequences (%)		
			PNA	TAReuk[a]	EKeuk[b]
Eukaryota	Eukaryota	56,829	11.8	62.7	35.4
	Fungi	9373	63.5	25.9	67.1
	Dikarya	7696	76.1	17.5	68.8
	Ascomycota	5156	81.3	26.0	71.7
Fungi	Basidiomycota	2538	65.4	0.2	62.7
	Chytridiomycota	319	6.6	71.8	70.8
	Cryptomycota	213	16.0	87.8	55.4
	Mucoromycota	752	4.3	60.0	60.9
	SAR	15,220	2.4	76.1	49.5
	SAR; Rhizaria	2522	1.2	64.5	74.9
	SAR; Rhizaria; Cercozoa	1374	0.1	70.4	74.5
	Chlorophyta	1908	1.7	85.7	76.6
	SAR; Alveolata	9140	2.6	76.3	28.9
	SAR; Alvelolata; Apicomplexa	1206	5.2	69.7	54.4
	SAR; Alveolata; Ciliophora	2935	3.6	71.9	58.0
	SAR; Stramenopiles	3556	2.8	83.7	40.9
	SAR; Stramenopiles; Ochrophyta	2392	2.0	88.1	86.5
Protists	Amoebozoa	1148	2.5	71.0	51.0
	Amoebozoa; Conosa; Archamoebae	15	0.0	66.7	6.7
	Amoebozoa; Lobosa; Discosea	431	4.9	80.7	71.2
	Amoebozoa; Lobosa; Tubulinea	132	0.0	80.3	40.9
	Discoba	1032	0.2	2.4	4.0
	Pseudofungi; Hyphochytridiomycetes	7	0.0	100.0	85.7
	Sagenista; Labyrinthulomycetes	392	1.5	78.1	73.0
Metazoa		21,032	0.5	69.2	1.5
Streptophyta		4538	0.1	62.4	64.0

[a] Stoeck et al., 2010.
[b] Simon et al., 2015 and Bower et al., 2004.

from soil DNA.

The TestPrime function of the SILVA data base showed that TAReuk perfectly matched with 25.9 % of all fungal sequences, more specifically 26.0 % of Ascomycota and only 0.2 % of Basidiomycota. In contrast, these primers showed a 76.1 % match with the SAR supergroup and a 71.0 % match with Amoebozoa. Furthermore, they also matched at similar efficiency with Metazoa and Streptophyta sequences. Compared to TAReuk, EKeuk revealed a clearly higher match with fungal sequences (67.1 % as compared to 25.9 %) for Ascomycota and Basidiomycota. With EKeuk, matches for the SAR supergroup were lower, overall with 49.5 %, but for Rhizaria (Cercozoa), specificities were similar to those seen with TAReuk. On the other hand, EKeuk was clearly less specific for SAR (Stramenopiles and Alvelolata) as well as Amoebozoa. For 18S rRNA genes from Metazoa, matches were much lower, with only 1.5 % as compared to 69.2 % with TAReuk, but for Streptophyta they were very similar. Overall, the *in silico* evaluation of two primer pairs revealed their potential for amplification of different spectra of soil eukaryotes with a pronounced overall preference of the TAReuk primer pair for the SAR supergroup, especially Alveolata, and also for Metazoa sequences, while EKeuk primers showed enhanced potential for the detection of the Ascomycota and Basidiomycota.

3.2. Evaluation of the antifungal PNA probe with two common primer pairs

3.2.1. Case study 1: "Land use"

For cropland but not forest soil, the presence of PNA along with TAReuk increased total reads of Eukaryotes and Protists, but reduced read counts of Fungi(Fig. S1). There was no such effect with EKeuk. TAReuk detected a relative higher abundance of 18S rRNA gene sequences from Metazoa and less from fungi, thus confirming their distinct *in silico* indicated specificities (Fig. 1A.). Good's coverage ranged from 0.68 to 0.95 and was neither affected by the primer pairs nor by presence of the PNA.

For cropland soil DNA and TAReuk, the addition of PNA decreased the relative abundance of fungal sequences, as expected by its intended design. The decrease in fungal sequences was, however, not accompanied by an increase of protists, but by Metazoa and the category "other sequences". From forest soil, unexpectedly, the presence of PNA had an opposite effect, increasing the proportion of fungal while reducing both metazoan and protistan 18S rRNA gene amplicon sequences.

In the presence of PNA, TAReuk primers increased fungal diversity from the cropland DNA, but decreased it in the paired forest soil (Table 2). With EKeuk, PNA decreased diversity and richness of both fungi and protists from cropland soil DNA, but no effect was tangible with forest soil DNA. For fungi, TAReuk detected a similar relative abundance of Ascomycota in both cropland and adjacent forest soils, as well as a similar abundance of Basidiomycota. In contrast, EKeuk indicated stronger prevalence of Ascomycota in cropland and Basidiomycota in forest soil (Fig. 1 B.). With TAReuk and cropland DNA, PNA had no effect on the proportion of ASV from Ascomycota but reduced the proportion of Basidiomycota and, in parallel, increased ASV of Mucoromycota. But with DNA from forest soil, PNA strongly increased the retrieval of ASV of Basidiomycota while decreasing Ascomycota. With EKeuk, PNA did not change results on the relative abundance of sequences from fungal phyla from cropland soil, but with DNA from forest soil, it reduced Basidiomycota and increased Ascomycota.

Protist derived 18S rRNA gene amplicons from TAReuk indicated that, for both cropland and forest soils, the most prevalent group was Cercozoa (Fig. 1 C.). A similar relative abundance of Cercozoa was indicated with EKeuk. With EKeuk, however, the most abundant group detected from forest DNA was Apicomplexa, which was a relatively small group in cropland samples. For TAReuk, irrespective of land use, the presence of PNA increased amplicon sequences from Conosa and reduced those from Chlorophyta. PNA also increased the retrieval of amplicons from Apicomplexa. For EKeuk, there was no detectable effect of PNA on assessing the relative abundance of protist sequences.

Principal component analyses (PCA) and PERMANOVA indicated for TAReuk significant differences in protist community compositions caused by the presence of the PNA (Fig. 2 A.). In combination with EKeuk, PNA effects were not tangible (Fig. 2 B.). However, soil fungal and protistan communities were significantly shaped by the land use with EKeuk, not TAReuk, which support the finding that EKeuk might reflect the actual community composition better than TAReuk.

3.2.2. Case study 2: "Cropland soils with variation in texture and tillage"

For the clay soil (CS) but not the other two soils, the presence of PNA increased with both primers the total read counts of Eukaryotes and Protists (Fig. S2). Good's coverage, ranged from 0.70 to 0.85, and there was no clear effect by primer pairs or presence of PNA (Table S1).

TAReuk retrieved from all three soils a high proportion of sequences from Metazoa, which did not occur at substantial amounts in EKeuk amplicon libraries (Fig. 3 A.). With EKeuk, the relative abundance of fungal sequences was much higher. Protist sequences were produced with both primer pairs at similar proportions. The presence of PNA did not increase the relative abundance detected for protists with any of the three soils, irrespective of the applied primer pairs.

For the three soils, with TAReuk, only the LS site showed a significant effect of PNA by increasing the diversity and richness of the amplicons from fungi and protists (Table 2). There was no comparable effect with the EKeuk primers.

For fungal community composition, most 18S rRNA gene amplicons obtained with EKeuk and TAReuk with soil DNA were affiliated with

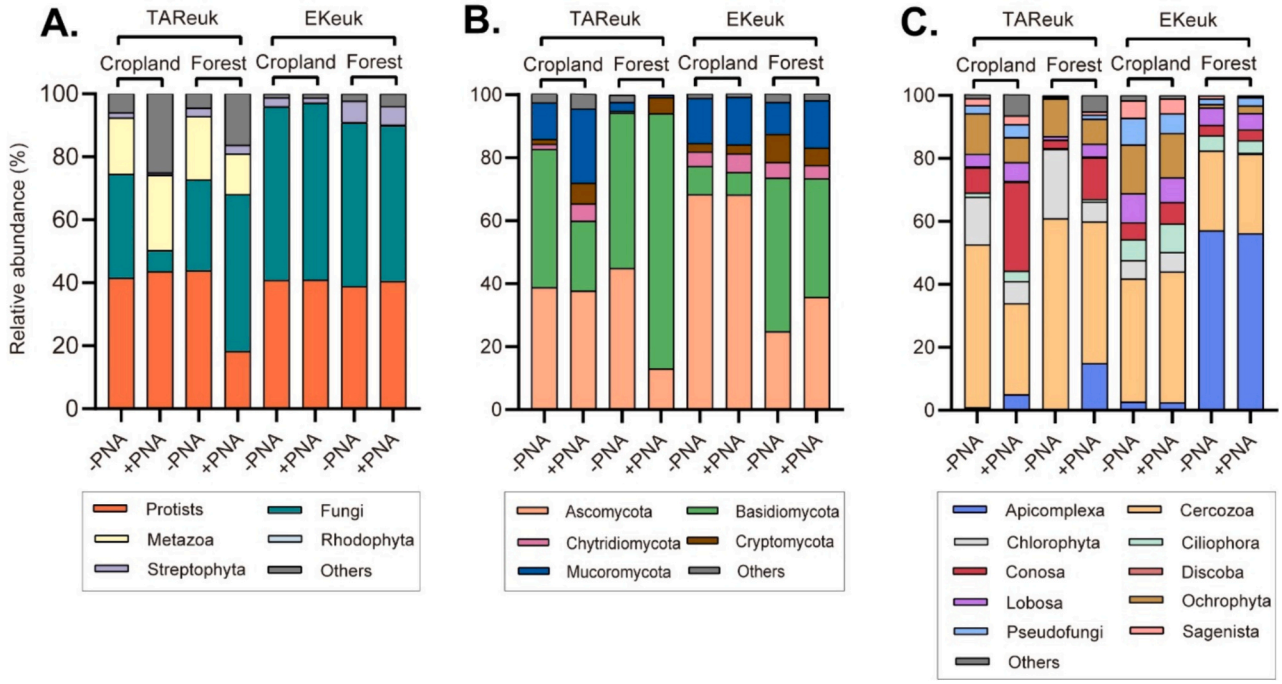


Fig. 1. Effect of PCR primers and the addition of an antifungal PNA on the relative abundance of 18S rRNA (variable region V4) gene amplicons (A.), and therein for fungal phyla (B.) and protist phyla (C.) obtained from cropland and forest DNA.

Table 2
Effect of primer pairs TAReuk and EKeuk, amplifying different variable regions of the 18S rRNA genes from soil DNA, in absence and presence of the PNA probe (+ PNA; control – PNA) on the Shannon diversity (H) and total richness (S) of amplicon sequence variants (ASV). Significant differences caused by the presence of PNA in each sample were investigated by Duncan's multiple comparison test. Significant values are indicated in bold and with *.

Primer pair		TAReuk				EKeuk			
ASV affiliation		Fungi		Protist		Fungi		Protist	
Soil type	PNA	H	S	H	S	H	S	H	S
Case study 1: Land use									
Forest	–	2.64	28	2.82	70	3.68	83	3.67	77
	+	**1.69***	30	**3.73***	120	3.73	74	3.64	73
Cropland	–	2.84	42	3.73	162	3.39	65	4.21	101
	+	**3.60***	84	**4.72***	371	**2.88***	**40***	**3.64***	**62***
Case study 2: Cropland soil with variation in texture and tillage									
Clay S	–	3.45	76	4.4	261	3.2	43	3.72	65
	+	3.55	80	4.36	275	3.22	48	3.92	79
Loam S	–	3.31	63	4	214	3.47	60	4.01	82
	+	3.68	**86***	**4.88***	**322***	3.36	54	3.77	68
Loam V	–	3.72	86	5	322	3.53	61	3.78	68
	+	3.76	91	5	316	3.49	57	3.68	60

Ascomycota, followed by approximately equal relative abundances of Basidiomycota and Mucoromycota (Fig. 3 B.). Variation between soils was minor. With EKeuk as compared to TAReuk, the proportion of ASV from Ascomycota was higher and from Basidiomycota lower. Irrespective of the primers, the PNA probe had no significant effect on the relative abundance of the fungal phyla.

For protists and for both primers, variation between the three soils was higher as compared to fungi (Fig. 3 C.). Cercozoa constituted the group with the most ASV, except for the clay soil (CS) with TAReuk, where Conosa dominated (Fig. 3 C). Sagenista were preferably detected with TAReuk and Ciliophora with EKeuk, thus showing alternative spectra. With PNA in combination with TAReuk, the relative abundance of Conosa increased, while in parallel Cercozoa declined. Other protist groups were less affected by PNA. With EKeuk no significant effect of PNA could be detected at any of the three soils.

PCA and PERMANOVA indicated that the PNA had no specific effect on the overall community compositions, neither with TAReuk nor with EKeuk and neither for fungi nor for protists (Fig. 4). Both primers showed that the main factor influencing fungal and protistan communities were the cropland sites (with variation in texture and tillage).

3.3. Amplicon library compositions and backwards testing the functionality of the PNA probe

From both case studies combined, the TAReuk primers yielded 1944 distinct eukaryotic ASV of which 1024 ASV were assigned to protists, 278 ASV to fungi, 325 ASV to Metazoa, and 310 ASV as “other”. The use of EKeuk resulted in 2780 distinct eukaryotic ASV of which 1456 ASV were protists, 1025 ASV fungi, 15 ASV Metazoa, and 284 ASV as “other”. In absence of the PNA probe, 192 fungal ASV amplified with EKeuk showed a perfect match with the DNA sequence of the PNA probe, while in presence of the PNA, no matching sequence was detected (Table 3),

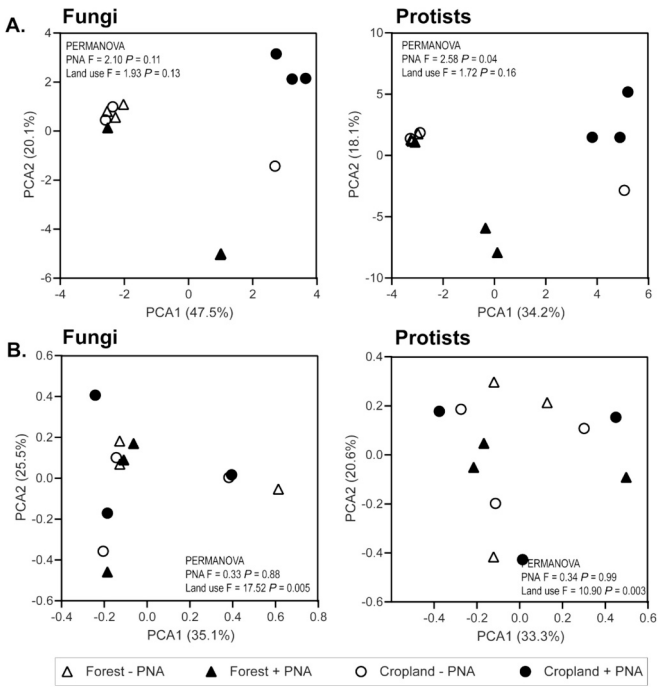


Fig. 2. Compositional differences of 18S rRNA V4 gene amplicons from Fungi and Protists obtained from cropland and forest soil DNA using primer pairs TAReuk (A.) and EKeuk (B.), visualized by principal component analysis (PCA) and PERMANOVA test.

thus indicating a 100 % efficiency of our probe in blocking PCR amplification of the fungal sequences. Surprisingly, with the TAReuk primers, no PNA complementary fungal sequences were detected, neither in presence nor in absence of the PNA probe, indicating that those primers did not only amplify less fungal sequences (278 ASV as compared to 1025 ASV) but also sequences with no 100 % match to the PNA antifungal probe.

4. Discussion

Both primer pairs of this study amplified the same variable region (V4) of the universal eukaryotic 18S rRNA genes from soil DNA, but with different preferences. EKeuk showed in contrast to TAReuk a very low specificity to amplify 18S rRNA sequences from Metazoa. Accordingly, almost no metazoan sequences were detected with EKeuk, neither from forest nor from the cropland soils, while with TAReuk, the relative abundance of such sequences was in the range of 15 to almost 50 %. Thus, the primer system EKeuk, originally tested with freshwater samples (Simon et al., 2015a; Simon et al., 2015b), was equally efficient for excluding metazoan sequences with soil DNA.

For the fungal community composition, the *in silico* analyses indicated that the specificity of EKeuk for amplifying Ascomycota and Basidiomycota was much higher than with TAReuk. Especially for Basidiomycota, TAReuk had almost no specificity at all, while other fungal groups *i.e.* Mucoromycota and Chytridiomycota were amplified by both primer pairs with comparable preference. In contrast to these *in silico* predictions, amplicon libraries with TAReuk contained usually 10 to 20 % taxa (ASV) assigned to Basidiomycota and this relative abundance tended to be even higher than with EKeuk. The amplification of Basidiomycota with such low specificity primers can be explained by the fact that the primers were not 100 % exclusive and that the soils may have harbored fungi with sequences not yet included in the gene databases. Since PCR-based approaches with "universal primers" depend on the quality of gene databases, this underlines the current limitations and need to collect more sequences from eukaryotic microorganisms. This can be obtained with alternative methods of collecting DNA sequences, *e.g.*, obtaining newly cultured isolates or sequencing soil metagenomes (Jansson and Hofmockel, 2018; Knight et al., 2018; Nilsson et al., 2019). It should be noted the *in silico* evaluation as applied in this study was based on 100 % primer match analyses, while for a successful amplification of sequences from soil DNA, primers binding with <100 % match probably also generated 18S rRNA gene amplicons (Parada et al., 2016). Furthermore, some differences between the *in silico* predictions and the actual amplicons retrieved may also have been caused by the fact that our study used two consecutive PCR steps (a two-step approach). The second PCR step, however, enhances a further amplification of already prevalent sequences while amplifying rare sequences with less efficiency, thus reducing the overall diversity (Finn et al., 2022).

The amplicon libraries with EKeuk indicated a higher prevalence of Ascomycota in cropland and of Basidiomycota in forest soils, while with TAReuk such differences were not detected. The results with EKeuk correspond to fungal diversity retrieved with ITS1 primers from cropland and forest sites of this region and therefore probably reflect the actual community composition better than TAReuk (Finn et al., 2023). In studies with arctic tundra soil, amplicon libraries with the TAReuk primers only contained a relative abundance of approx. 1 % fungal sequences (Almela et al., 2023), which appears to be an underestimate considering the expected general prevalence of fungi in soils, as assessed

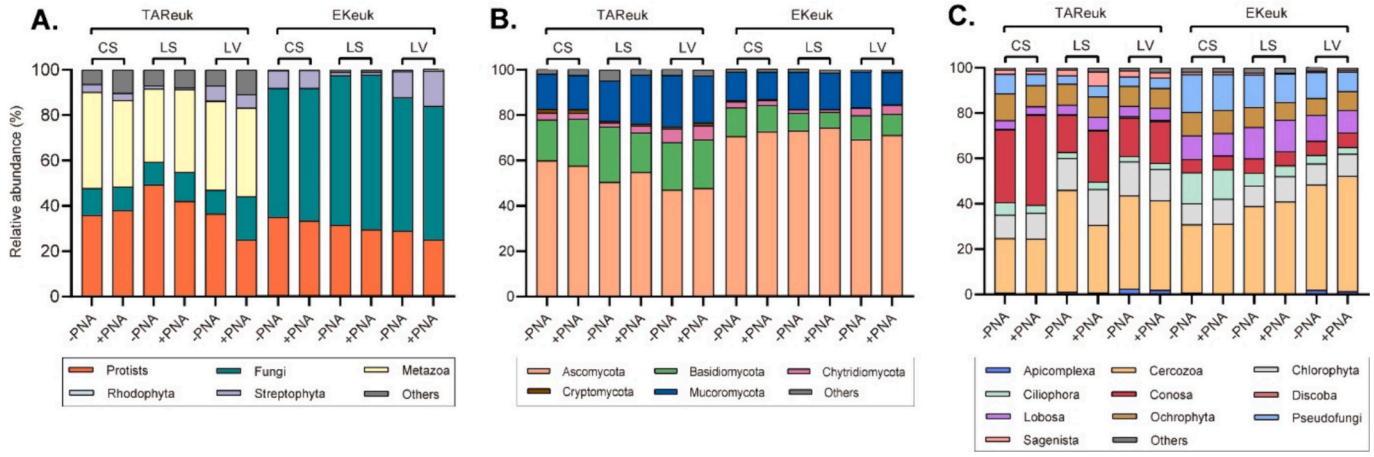


Fig. 3. Effect of the TAReuk and EKeuk primer pairs on the relative abundance of 18S rRNA V4 region gene amplicons obtained from cropland soil DNA. A. Relative abundance of all quality filtered amplicons (A.), of fungal (B.) and of protist (C.) assigned amplicons. Soil samples originate from three neighboring fields, *i.e.*, CS, clay soil, conservative tillage; LS, loam soil, conservative tillage; LV, loam soil, conventional tillage.

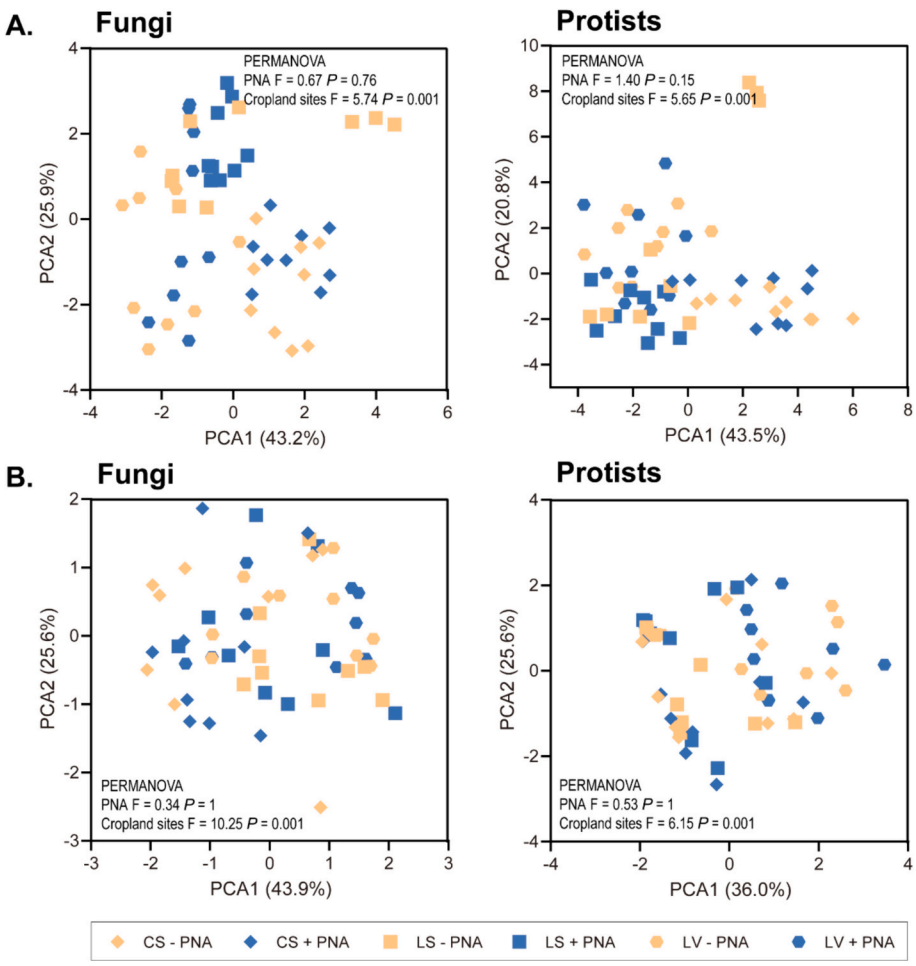


Fig. 4. Principal component analyses (PCA) and PERMANOVA test to visualize the compositional differences of the 18S rRNA V4 gene amplicon sequence libraries obtained from cropland soil DNA with primers TAReuk (A.) and EKeuk (B.) in presence (+) and absence (−) of the antifungal PNA probe of this study. CS, clay soil, conservation tillage; LS, loam soil, conservative tillage; LV, loam soil, conventional tillage.

Table 3
Prevalence of DNA amplicon sequences retrieved in this study with complementary sequences to the PNA probe in its absence (− PNA) and presence (+ PNA) during PCR amplifying the fungal V4 18S rRNA gene region.

Primer system	Total fungal ASV		Fungal ASV with complementary PNA sequences	
	- PNA	+ PNA	- PNA	+ PNA
TAReuk	266	269	0	0
EKeuk	697	617	192	0

with different PCR-independent methods (Rousk et al., 2010a; Rousk et al., 2010b; Tveit, 2012).To avoid the anti-protist bias, other studies with agricultural soil utilized in addition to a fungi-specific primer pair targeting the V7 region (Borneman and Hartin, 2000) TAReuk for focusing on protists (Du et al., 2022; Zhao et al., 2019).

For protists, the *in silico* analyses indicated different primer specificities of EKeuk and TAReuk especially for Conosa (Amoebozoa), with 10-fold higher specificity for the latter. Members of this group occurred in all soil DNA samples and for the three cropland soils (Case study 2), their relative abundance was about two-fold higher with TAReuk than with EKeuk. Differences of primer specificities, as indicated by the *in silico* analyses, were less pronounced for the other protist groups, and, in fact, both detected the same higher taxonomic ranks in both case studies. For the land use Case study 1, the TAReuk primers indicated that Cercozoa were the dominant group in cropland and forest soil, and the dominance in cropland soils was confirmed by EKeuk. However, in forest soil, Apicomplexa showed a higher relative abundance, which was surprising because primer specificities were slightly higher with TAReuk. And, in fact, with the cropland soils of Case study 2, the relative abundance of Apicomplexa was higher with EKeuk compared to TAReuk. Inconsistencies between predicted and detected sequence diversity can be explained by the comparably scarce DNA sequences pf Apicomplexa and other groups in the gene databases. This also related to Lobosa, which were detected at the cropland field sites with higher relative abundance with EKeuk than with TAReuk, even though *in silico* data suggested that TAReuk would have a higher specificity for obtaining ASV from this group. Overall, the strong differences between *in silico* specificities and the actual phylogenetic identity of ASV amplified from soil DNA indicates the current limitation of using *in silico* predictions for PCR amplifying specific phylogenetic groups and taxa from complex highly diverse metagenomes, *i.e.* in soil (Tedersoo et al., 2015). Probably the major barriers are linked to the comparably small database for protist DNA sequences (Geisen et al., 2023) and variation in primer binding efficiencies below 100 % sequence match (Parada et al., 2016).

For our antifungal PNA probe, it was expected that it would reduce the amplification of fungal sequences and consequently increase the relative abundance of ASV from other eukaryotes, including protists. It was also expected that PNA would more strongly affect the PCR amplification with the EKeuk primer pair which favored Basidiomycota. Surprisingly however, the presence of PNA did not reduce the relative

abundance of fungi as compared to protists or other sequences with EKeuk, neither with the different cropland soil DNA of both case studies, nor with forest soil DNA. Apparently, there were still sufficient amplifiable fungal sequences in soil, which could not be reduced with the PNA because they did not match the representative sequences in the Silva database, as confirmed by amplicon sequencing. In Case study 1, with TAReuk primers, the PNA probe effectively reduced with cropland soil DNA the proportion of fungi in favor of protists, just as intended by our design, but this must have been an indirect effect, considering that amplicon sequencing indicated that all fungal sequences amplified with TAReuk did not have a 100 % match with the PNA. In fact, with forest soil DNA, the fungal proportion increased in presence of the PNA. Whether land use is a factor influencing these outcomes cannot be concluded because only one single paired land use site was analyzed in this study. Larger scale land-use studies indicated consistently specific differences between cropland and forest soils in the microbial community compositions (Szoboszlay et al., 2017).

While for both case studies the PNA probes did not cause a specific inhibition of fungal sequences as compared to protists or other eukaryotes, there were effects of PNA with the TAReuk on the diversity of protist groups retrieved from soil DNA. In the cropland and forest soils used for Case study 1, the presence of PNA increased the detection of Conosa and Apicomplexa while it reduced the relative abundance of Chlorophyta. In the cropland soils of the Case study 2 study, there was also an increased detection of Conosa in presence of the PNA, accompanied by reduction of the relative abundance of Cercozoa. In contrast, PNA had no such effect with the EKeuk primers. Conosa are a group within the Amoebozoa (Cavalier-Smith et al., 2015) and the number of sequences accessible in the databases was very low with $n = 15$, as compared to 1148 sequences of the higher rank Amoebozoa (Guillou et al., 2013). As suggested by their detection with both primer systems of this study, the contribution of protists from this group Conosa in soil, however is not minor, ranging across all samples between 3 and 40 % in relative abundance. Conosa encompass slime molds (Myxomycota), which are typically widespread and prevalent in soil, where they participate as saprophytes in the decomposition of organic substrates *i.e.* plant litter (Asiloglu et al., 2021; Cainelli et al., 2020). The discrepancy between the few 18S rRNA gene sequences in the databases and their apparent prevalence in soil is unlikely due to their low diversity but more probably linked to the fact that their larger 18S rRNA gene amplicons are not amplified with equal efficiency from soil DNA as shorter versions. Consequently, slime molds and their relatives are an understudied group not yet sufficiently represented in the DNA sequence databases (Stephenson et al., 2011). The antifungal PNA probe of this study in combination with the TAReuk primer system could be a useful tool to study the diversity and ecology of this group in soil.

Overall, our study demonstrates that the spectrum of 18S rRNA genes amplified from soil DNA can be greatly enhanced by using different primers targeting the same variable region. While the antifungal PNA probe developed in this study indicated high specificity by *in silico* analyses, its performance to suppress fungal sequences in favor of enhancing protist and metazoan sequences with soil DNA was strongly limited to specific soil DNA samples and also showed unanticipated differences between land use (cropland *vs.* forest). These limitations can most likely be attributed to (i) primer binding to <100 % complementary sequences, (ii) the two step PCR protocol suppressing amplification of rare sequences, and (iii) that the diversity of sequences from microbial eukaryotes in soil is still much higher than currently reflected in the DNA sequence databases. With more sequences enriching the current databases, the design of more effective PNA, or using a combination of different PNA probes in conjunction with different primers could provide a promising approach for unravelling the so far hidden diversity of soil microbial eukaryotes. The PNA of this study demonstrates in combination with the TAReuk primer its potential to enhance the detection of Conosa in cropland soils and, thus, improve our understanding of the diversity and ecology of this highly abundant, yet understudied group in soil.

Funding

The study was founded by the National Monitoring of Biodiversity in Agricultural Landscapes (MonViA) project from the German Federal Ministry of Food and Agriculture.

CRediT authorship contribution statement

Haotian Wang: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Kenneth Dumack:** Writing – review & editing, Resources, Methodology. **Daniel V. Rissi:** Methodology, Investigation. **Damien R. Finn:** Writing – review & editing, Software, Data curation, Conceptualization. **Michael Bonkowski:** Writing – review & editing, Methodology, Conceptualization. **Christoph C. Tebbe:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Karin Trescher, Jana Usarek and Britta Müller contributed to this work with their excellent technical support, which we gratefully acknowledge. We thank Julia Schroeder and Joachim Brunotte for providing soil samples.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apsoil.2024.105464.

References

Adl, S.M., Bass, D., Lane, C.E., Lukes, J., Schoch, C.L., Smirnov, A., Agatha, S., Berney, C., Brown, M.W., Burki, F., Cardenas, P., Cepicka, I., Chistyakova, L., del Campo, J., Dunthorn, M., Edvardsen, B., Eglit, Y., Guillou, L., Hampl, V., Heiss, A.A., Hoppenrath, M., James, T.Y., Karnkowska, A., Karpov, S., Kim, E., Kolisko, M., Kudryavtsev, A., Lahr, D.J.G., Lara, E., Le Gall, L., Lynn, D.H., Mann, D.G., Massana, R., Mitchell, E.A.D., Morrow, C., Park, J.S., Pawlowski, J.W., Powell, M.J., Richter, D.J., Rueckert, S., Shadwick, L., Shimano, S., Spiegel, F.W., Torruella, G., Youssef, N., Zlatogursky, V., Zhang, Q.Q., 2019. Revisions to the classification, nomenclature, and diversity of eukaryotes. J. Eukaryot. Microbiol. 66, 4–119.

Almela, P., Rico, E., Velazquez, D., Verleyen, E., Quesada, A., 2023. Soil moisture drives differences in the diversity and trophic complexity of high Arctic tundra soils. FEMS Microbiol. Ecol. 99.

Asiloglu, R., Samuel, S.O., Sevilir, B., Akca, M.O., Bozkurt, P.A., Suzuki, K., Murase, J., Turgay, O.C., Harada, N., 2021. Biochar affects taxonomic and functional community composition of protists. Biol. Fertil. Soils 57, 15–29.

Banos, S., Lentendu, G., Kopf, A., Wubet, T., Glockner, F.O., Reich, M., 2018. A comprehensive fungi-specific 18S rRNA gene sequence primer toolkit suited for diverse research issues and sequencing platforms. BMC Microbiol. 18.

Bell, J.K., Helgason, B., Siciliano, S.D., 2021. Brassica napus phyllosphere bacterial composition changes with growth stage. Plant Soil 464, 501–516.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37, 852–857.

Borneman, J., Hartin, R.J., 2000. PCR primers that amplify fungal rRNA genes from environmental samples. Appl. Environ. Microbiol. 66, 4356–4360.

Bower, S.M., Carnegie, R.B., Goh, B., Jones, S.R., Lowe, G.J., Mak, M.W., 2004. Preferential PCR amplification of parasitic protistan small subunit rDNA from metazoan tissues. J. Eukaryot. Microbiol. 51, 325–332.
Cainelli, R., de Haan, M., Meyer, M., Bonkowski, M., Fiore-Donno, A.M., 2020. Phylogeny of Physarida (Amoebozoa, Myxogastria) based on the small-subunit ribosomal RNA gene, redefinition of Physarum pusillum s. str. and reinstatement of P. gravidum Morgan. J. Eukaryot. Microbiol. 67, 327–336.
Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nat. Methods 13, 581–583.
Callahan, B.J., McMurdie, P.J., Holmes, S.P., 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 11, 2639–2643.
Catlett, D., Matson, P.G., Carlson, C.A., Wilbanks, E.G., Siegel, D.A., Iglesias-Rodriguez, M.D., 2020. Evaluation of accuracy and precision in an amplicon sequencing workflow for marine protist communities. Limnol. Oceanogr.-Methods 18, 20–40.
Cavalier-Smith, T., Fiore-Donno, A.M., Chao, E., Kudryavtsev, A., Berney, C., Snell, E.A., Lewis, R., 2015. Multigene phylogeny resolves deep branching of Amoebozoa. Mol. Phylogenet. Evol. 83, 293–304.
Choi, T.J., Malik, A., An, H.E., Kim, J.I., Do, T.D., Kim, C.B., 2022. Seasonal diversity of microeukaryotes in the Han River, Korea through 18S rRNA gene metabarcoding. Evol. Bioinforma. 18.
Du, S., Li, X.-Q., Hao, X., Hu, H.-W., Feng, J., Huang, Q., Liu, Y.-R., 2022. Stronger responses of soil protistan communities to legacy mercury pollution than bacterial and fungal communities in agricultural systems. ISME Commun. 2, 69.
Finn, D.R., Samad, M.S., Tebbe, C.C., 2022. One-step PCR amplicon sequencing libraries perform better than two-step when assessing soil microbial diversity and community profiles. FEMS Microbiol. Lett. 369, fnac079.
Finn, D.R., Schroeder, J., Samad, M.S., Poeplau, C., Tebbe, C.C., 2023. Importance of sample pre-treatments for the DNA-based characterization of microbiomes in cropland and forest soils. Soil Biol. Biochem. 184.
Fitzpatrick, C.R., Lu-Irving, P., Copeland, J., Guttman, D.S., Wang, P.W., Baltrus, D.A., Dlugosch, K.M., Johnson, M.T.J., 2018. Chloroplast sequence variation and the efficacy of peptide nucleic acids for blocking host amplification in plant microbiome studies. Microbiome 6.
Geisen, S., Bandow, C., Römbke, J., Bonkowski, M., 2014. Soil water availability strongly alters the community composition of sil protists. Pedobiologia 57, 205–213.
Geisen, S., Lara, E., Mitchell, E., 2023. Contemporary issues, current best practice and ways forward in soil protist ecology. Mol. Ecol. Resour. 23, 1477–1487.
Gloor, G.B., Macklaim, J.M., Fernandes, A.D., 2016. Displaying variation in large datasets: plotting a visual summary of effect sizes. J. Comput. Graph. Stat. 25, 971–979.
Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., 2012. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41, D597–D604.
Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., del Campo, J., Dolan, J.R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W., Lara, E., Le Bescot, N., Logares, R., Mahé, F., Massana, R., Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.L., Siano, R., Stoeck, T., Vaulot, D., Zimmermann, P., Christen, R., 2013. The Protist Ribosomal Reference database (PR²): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41, D597–D604.
Holzer, A.S., Wootten, R., Sommerville, C., 2007. The secondary structure of the unusually long 18S ribosomal RNA of the myxozoan Sphaerospora truttae and structural evolutionary trends in the Myxozoa. Int. J. Parasitol. 37, 1281–1295.
Hugerth, L.W., Muller, E.E.L., Hu, Y.O.O., Lebrun, L.A.M., Roume, H., Lundin, D., Wilmes, P., Andersson, A.F., 2014. Systematic design of 18S rRNA gene primers for determining eukaryotic diversity in microbial consortia. PLoS One 9.
Jackrel, S.L., Owens, S.M., Gilbert, J.A., Pfister, C.A., 2017. Identifying the plant-associated microbiome across aquatic and terrestrial environments: the effects of amplification method on taxa discovery. Mol. Ecol. Resour. 17, 931–942.
Jansson, J.K., Hofmockel, K.S., 2018. The soil microbiome - from metagenomics to metaphenomics. Curr. Opin. Microbiol. 43, 162–168.
Knight, R., Vrbanac, A., Taylor, B.C., Aksenov, A., Callewaert, C., Debelius, J., Gonzalez, A., Kosciolek, T., McCall, L.I., McDonald, D., Melnik, A.V., Morton, J.T., Navas, J., Quinn, R.A., Sanders, J.G., Swafford, A.D., Thompson, L.R., Tripathi, A., Xu, Z.J.Z., Zaneveld, J.R., Zhu, Q.Y., Caporaso, J.G., Dorrestein, P.C., 2018. Best practices for analysing microbiomes. Nat. Rev. Microbiol. 16, 410–422.
Laure, G., Dipankar, B., Stéphane, A., David, B., Cédric, B., Lucie, B., Christophe, B., Gaétan, B., Colomban, D.V., Johan, D., 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41, D597–D604.
Lefevre, E., Gardner, C.M., Gunsch, C.K., 2020. A novel PCR-clamping assay reducing plant host DNA amplification significantly improves prokaryotic endo-microbiome community characterization. FEMS Microbiol. Ecol. 96.
Lentendu, G., Wubet, T., Chatzinotas, A., Wilhelm, C., Buscot, F., Schlegel, M., 2014. Effects of long-term differential fertilization on eukaryotic microbial communities in an arable soil: a multiple barcoding approach. Mol. Ecol. 23, 3341–3355.
Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17.1.
Nielsen, P.E., 2004. PNA technology. Mol. Biotechnol. 26, 233–248.
Nilsson, R.H., Anslan, S., Bahram, M., Wurzbacher, C., Baldrian, P., Tedersoo, L., 2019. Mycobiome diversity: high-throughput sequencing and identification of fungi. Nat. Rev. Microbiol. 17, 95–109.
Oksanen, J., 2010. Vegan: community ecology package. http://vegan.r-forge.r-project.org/.
Oliverio, A.M., Geisen, S., Delgado-Baquerizo, M., Maestre, F.T., Turner, B.L., Fierer, N., 2020. The global-scale distributions of soil protists and their contributions to belowground systems. Sci. Adv. 6, eaax8787.
Parada, A.E., Needham, D.M., Fuhrman, J.A., 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ. Microbiol. 18, 1403–1414.
Pascoal, F., Costa, R., Magalhaes, C., 2021. The microbial rare biosphere: current concepts, methods and ecological principles. FEMS Microbiol. Ecol. 97.
Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S.S., Cepicka, I., Decelle, J., Dunthorn, M., Fiore-Donno, A.M., Gile, G.H., Holzmann, M., Jahn, R., Jirku, M., Keeling, P.J., Kostka, M., Kudryavtsev, A., Lara, E., Lukes, J., Mann, D.G., Mitchell, E.A.D., Nitsche, F., Romeralo, M., Saunders, G.W., Simpson, A.G.B., Smirnov, A.V., Spouge, J.L., Stern, R.F., Stoeck, T., Zimmermann, J., Schindel, D., de Vargas, C., 2012. CBOL Protist Working Group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. PLoS Biol. 10.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glockner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D141–D145.
Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahe, F., 2016. VSEARCH: a versatile open source tool for metagenomics. Peerj 4.
Rousk, J., Bååth, E., Brookes, P.C., Lauber, C.L., Lozupone, C., Caporaso, J.G., Knight, R., Fierer, N., 2010a. Soil bacterial and fungal communities across a pH gradient in an arable soil. ISME J. 4, 1340–1351.
Rousk, J., Brookes, P.C., Bååth, E., 2010b. Investigating the mechanisms for the opposing pH relationships of fungal and bacterial growth in soil. Soil Biol. Biochem. 42, 926–934.
Sambrook, J.R.D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.
Sievers, F., Higgins, D.G., 2014. Clustal Omega, accurate alignment of very large numbers of sequences. In: Multiple Sequence Alignment Methods. Springer, pp. 105–116.
Simon, M., Jardillier, L., Deschamps, P., Moreira, D., Restoux, G., Bertolino, P., Lopez-Garcia, P., 2015a. Complex communities of small protists and unexpected occurrence of typical marine lineages in shallow freshwater systems. Environ. Microbiol. 17, 3610–3627.
Simon, M., Lopez-Garcia, P., Deschamps, P., Moreira, D., Restoux, G., Bertolino, P., Jardillier, L., 2015b. Marked seasonality and high spatial variability of protist communities in shallow freshwater systems. ISME J. 9, 1941–1953.
Stephenson, S.L., Fiore-Donno, A.M., Schnittler, M., 2011. Myxomycetes in soil. Soil Biol. Biochem. 43, 2237–2242.
Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.W., Richards, T.A., 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Mol. Ecol. 19, 21–31.
Szoboszlay, M., Dohrmann, A.B., Poeplau, C., Don, A., Tebbe, C.C., 2017. Impact of land-use change and soil organic carbon quality on microbial diversity in soils across Europe. FEMS Microbiol. Ecol. 93, fix146.
Taerum, S.J., Steven, B., Gage, D.J., Triplett, L.R., 2020. Validation of a PNA clamping method for reducing host DNA amplification and increasing eukaryotic diversity in rhizosphere microbiome studies. Phytobiomes J. 4, 291–302.
Taib, N., Mangot, J.F., Domaizon, I., Bronner, G., Debroas, D., 2013. Phylogenetic affiliation of SSU rRNA genes generated by massively parallel sequencing: new insights into the freshwater protist diversity. PLoS One 8.
Tan, S.J., Liu, H.B., 2018. Unravel the hidden protistan diversity: application of blocking primers to suppress PCR amplification of metazoan DNA. Appl. Microbiol. Biotechnol. 102, 389–401.
Tedersoo, L., Bahram, M., Polme, S., Anslan, S., Riit, T., Koljalg, U., Nilsson, R.H., Hildebrand, F., Abarenkov, K., 2015. Response to comment on "global diversity and geography of soil fungi". Science 349.
Tveit, M.M.N., 2012. Interaction Studies of the Fusarium-wheat System: Fungal-fungal and Plant-pathogen Interactions. Norwegian University of Life Sciences, Ås.
Vaulot, D., Geisen, S., Mahe, F., Bass, D., 2022. pr2-primers: An 18S rRNA primer database for protists. Mol. Ecol. Resour. 22, 168–179.
von Wintzingerode, F., Landt, O., Ehrlich, A., Gobel, U.B., 2000. Peptide nucleic acid-mediated PCR clamping as a useful supplement in the determination of microbial diversity. Appl. Environ. Microbiol. 66, 549–557.
Xie, Q., Lin, J.Z., Qin, Y., Zhou, J.F., Bu, W.J., 2011. Structural diversity of eukaryotic 18S rRNA and its impact on alignment and phylogenetic reconstruction. Protein Cell 2, 161–170.
Yarza, P., Yilmaz, P., Panzer, K., Glockner, F.O., Reich, M., 2017. A phylogenetic framework for the kingdom Fungi based on 18S rRNA gene sequences. Mar. Genomics 36, 33–39.

Yu, Y.J., Yang, Q.Y., Petropoulos, E., Zhu, T.B., 2022. ITS3/ITS4 outperforms other ITS region and 18S rRNA gene primer sets for amplicon sequencing of soil fungi. Eur. J. Soil Sci. 73.

Yuan, X.-L., Cao, M., Liu, X.-M., Du, Y.-M., Shen, G.-M., Zhang, Z.-F., Li, J.-H., Zhang, P., 2018. Composition and genetic diversity of the Nicotiana tabacum microbiome in different topographic areas and growth periods. Int. J. Mol. Sci. 19, 3421.

Zhao, Z.B., He, J.Z., Geisen, S., Han, L.L., Wang, J.T., Shen, J.P., Wei, W.X., Fang, Y.T., Li, P.P., Zhang, L.M., 2019. Protist communities are more sensitive to nitrogen fertilization than other microorganisms in diverse agricultural soils. Microbiome 7.