Journal of Microbiological Methods 125 (2016) 91–97

Contents lists available at ScienceDirect

Journal of Microbiological Methods

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

Novel oligonucleotide primers reveal a high diversity of microbes which drive phosphorous turnover in soil

Fabian Bergkemper [a], Susanne Kublik [a], Friederike Lang [b], Jaane Krüger [b], Gisle Vestergaard [a], Michael Schloter [a,*], Stefanie Schulz [a]

[a] Research Unit Environmental Genomics, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany
[b] Professur für Bodenökologie, Albert-Ludwigs-Universität Freiburg, Bertoldstr. 17, 79085 Freiburg i. Br., Germany

ARTICLE INFO

Article history:
Received 29 March 2016
Received in revised form 16 April 2016
Accepted 16 April 2016
Available online 19 April 2016

Keywords:
Phosphorus turnover
Forest soil
phoD
phoN
appA
pitA
pstS

ABSTRACT

Phosphorus (P) is of central importance for cellular life but likewise a limiting macronutrient in numerous environments. Certainly microorganisms have proven their ability to increase the phosphorus bioavailability by mineralization of organic-P and solubilization of inorganic-P. On the other hand they efficiently take up P and compete with other biota for phosphorus. However the actual microbial community that is associated to the turnover of this crucial macronutrient in different ecosystems remains largely anonymous especially taking effects of seasonality and spatial heterogeneity into account. In this study seven oligonucleotide primers are presented which target genes coding for microbial acid and alkaline phosphatases (*phoN*, *phoD*), phytases (*appA*), phosphonatases (*phnX*) as well as the quinoprotein glucose dehydrogenase (*gcd*) and different P transporters (*pitA*, *pstS*). Illumina amplicon sequencing of soil genomic DNA underlined the high rate of primer specificity towards the respective target gene which usually ranged between 98% and 100% (*phoN*: 87%). As expected the primers amplified genes from a broad diversity of distinct microorganisms. Using DNA from a beech dominated forest soil, the highest microbial diversity was detected for the alkaline phosphatase (*phoD*) gene which was amplified from 15 distinct phyla respectively 81 families. Noteworthy the primers also allowed amplification of *phoD* from 6 fungal orders. The genes coding for acid phosphatase (*phoN*) and the quinoprotein glucose dehydrogenase (*gcd*) were amplified from 20 respectively 17 different microbial orders. In comparison the phytase and phosphonatase (*appA*, *phnX*) primers covered 13 bacterial orders from 2 different phyla respectively. Although the amplified microbial diversity was apparently limited both primers reliably detected all orders that contributed to the P turnover in the investigated soil as revealed by a previous metagenomic approach. Genes that code for microbial P transporter (*pitA*, *pstS*) were amplified from 13 respectively 9 distinct microbial orders. Accordingly the introduced primers represent a valuable tool for further analysis of the microbial community involved in the turnover of phosphorus in soils but most likely also in other environments.

1. Introduction

Phosphorus (P) is an essential macronutrient for all biota on earth as it is integral for processes of cellular bioenergetics, the formation of lipid bilayers and the genetic backup (Elser, 2012). In most natural ecosystems in addition to nitrogen (N), P is a major growth limiting factor of primary production (Vitousek et al., 2010). Although many soils contain extensive stocks of total P, the bioavailability of soluble orthophosphate, which can be used by most biota is low (Rodriguez and Fraga, 1999). Consequently P is considered as the most inaccessible and unavailable of all soil nutrients (Holford, 1997). Therefore the prominent role of microorganisms for the turnover of soil phosphorus is generally accepted, since they can increase the P availability by different means, which lead to improved P nutrition of plants and other biota (Richardson and Simpson, 2011). Especially the metabolic traits which perform the mineralization of organic-P and the solubilization of inorganic-P are of peculiar interest. Most previous studies mainly focused on the characterization of the respective enzymes using cultivated bacterial and fungal strains (Rodriguez and Fraga, 1999). However only few attempts have been made to directly target genes that drive the turnover of soil P in microbial communities derived from environmental samples without introducing the bias caused by isolation. This might be of great relevance since strains that perform well under controlled conditions might easily be outcompeted in natural environments (Rodriguez and Fraga, 1999). It was not until 2008 that Sakurai et al. (2008) developed PCR primers specifically targeting the alkaline phosphatase *phoD* gene. Since an amplification bias towards Alphaproteobacteria was recently observed (Tan et al., 2013), Ragot et al. (2015) introduced a new set of primers which increased the covered diversity of *phoD* genes in soils by the factor of 7. For genes encoding other key processes of the

* Corresponding author at: Research Unit Environmental Genomics, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764, Germany.
E-mail address: schloter@helmholtz-muenchen.de (M. Schloter).

http://dx.doi.org/10.1016/j.mimet.2016.04.011
0167-7012/

microbial P turnover like the mineralization of organic-P, the solubilization of inorganic-P or phosphorus uptake, primer systems are still missing.

Recently, we performed a metagenomic analysis based on whole genome shotgun (WGS) sequencing in two undistorted beech forest soils to gain insights into the microbial P turnover in soil (Bergkemper et al., 2015). The results highlighted Rhizobiales, Actinomycetales and Acidobacteriales as drivers for P turnover but also rare orders like Solibacterales contributed to the turnover of soil P. Dominating processes were the uptake of P by the phosphate inorganic transporter (Pit) and the phosphate specific transport (Pst) system, the solubilization of inorganic P and the mineralization of organic P by alkaline and acid phosphatases, phosphonatases and phytases. In contrast, Glycerol-3-phosphate transporter or C–P lyases were of minor importance. As the obtained sequencing depth of metagenomic approaches is still far away from allowing quantitative conclusions (Delmont et al., 2011), metagenomic data can serve as a starting point for a targeted primer development to answer questions about spatial and temporal distribution of microbial key players (Schöler et al., 2016). Moreover due to the important role of microorganisms for P turnover we propose that there is a need to investigate the entire set of enzymes (genes) involved in processes of microbial P mineralization, solubilization and uptake to better understand the turnover of this crucial nutrient in soil. Thus it was the aim of this study to develop primer systems suitable for high-throughput amplicon sequencing as well as quantitative real-time PCR approaches. Based on the obtained metagenomic data we chose 7 marker genes, which were highly dominant in the metagenomes of the two forest soils. Overall we covered important steps of the solubilization of inorganic P, the mineralization of organic P and the cellular P uptake. For targeting P solubilization processes, which are mainly attributed to the efflux of protons and organic anions during the oxidation of glucose and other aldose sugars (Goldstein, 1994), we targeted the quinoprotein glucose dehydrogenase (*gcd*) gene (Cleton-Jansen et al., 1990). With respect to the mineralization of organic-P three different classes of enzymes were investigated: (i) Nonspecific acid phosphohydrolases (NSAPs) and alkaline phosphatases (ALPs) perform the dephosphorylation of phosphoester and — anhydride bonds. Here we focused on the NSAP class A (*phoN*) (Rossolini et al., 1998) and the ALP PhoD (*phoD*) (Eder et al., 1996). (ii) The mineralization of more complex myo-Inositol-1,2,3,4,5,6-hexakisphosphates (IP_6) is catalyzed by microbial phytases. Especially enzymes which are classified as 6-phytases (*appA*) (Golovan et al., 2000) were targeted here. (iii) The enzymatic cleavage of relatively stable carbon–phosphorus bonds, which occur in natural and synthetic organophosphonates, is performed by C–P lyases and phosphonoacetaldehyde hydrolases. The latter one (phosphonatase; *phnX*) (Hsieh and Wanner, 2010) was further investigated in this study. Moreover, microbes also compete for the available P with other biota as they have efficient phosphate uptake systems (Pst, Pit transporter). Therefore, we also targeted the key genes *pitA* and *pstS* (Hsieh and Wanner, 2010).

We aimed to cover a broad diversity of distinct microorganisms for the mentioned processes. The primer specificity towards the individual target genes and the diversity of the amplified microbial communities were investigated by Illumina amplicon sequencing of genomic DNA extracted from beech forest soil, which has also been included in the metagenomic analysis described above.

2. Material and methods

2.1. Site description and soil sampling

For this study soil samples were taken from a beech (*Fagus sylvatica*) dominated German forest site located in the Bavarian Rhoen Mountains near Bad Brueckenau (BBR) (50.352009°N, 9.929028°E). The stand has an average age of 120 years and is part of the International Co-operative Program for the Assessment and Monitoring of Air Pollution Effects on Forests (ICP Level II). The forest site reaches up to 850 m above sea level and the mean annual precipitation and temperature are 1031 mm respectively 5.8 °C. According to the World Reference Base (WRB) for Soil Resources the soil is classified as Dystric Skeletic Cambisol with Mull-like Moder and basalt as the substrate. The soil organic layer (Of-horizon) is characterized by a total carbon (C) content of 481.81 mg/g, a total nitrogen (N) content of 20.66 mg/g, a total phosphorus (P) content of 1.57 mg/g and a pH (H_2O) of 5.30 (pers. comm. F. Lang, Freiburg). Soil samples were taken from the organic layer in September 2015 after the removal of forest litter and immediately stored on dry ice.

2.2. Nucleic acid extraction

DNA was directly extracted from 0.5 g of frozen soil (−80 °C) using Lysing Matrix E tubes (MP Biomedicals, France) and Precellys 24 (Bertin Technologies, France) according to Töwe et al. (2011). Additionally a negative control of the extraction procedure was performed without soil input (NKE). Total genomic DNA was quantified using the Quant-iT PicoGreen kit (Life Technologies, USA) and stored at −20 °C for further analysis.

2.3. Primer development

Oligonucleotide primers were designed for microbial genes encoding enzymes that catalyze the mineralization of soil organic phosphorus (P) (*phoD*, *phoN*, *appA*, *phnX*), the solubilization of inorganic-P (*gcd*) as well as genes encoding microbial phosphate transport systems (*pitA*, *pstS*). The investigated genes together with the EC (TC) numbers of corresponding proteins are listed in Table 1. Oligonucleotide primers were designed in a way to amplify target genes from a broad diversity of different soil microorganisms. However to ensure a high level of specificity during PCR the grade of primer degeneracy was limited to a maximum of 64-fold. Since the amplicon size ranged from 147 bp to 375 bp the primers were suitable for both next generation sequencing (Illumina manual "16S Metagenomic Sequencing Library Preparation") as well as quantitative real-time PCR approaches (Karsai et al., 2002). Moreover the calculated annealing temperature of all primers was set to 60 °C. For primer development sequences of the investigated genes were derived from the NCBI Protein database (May 2015). The obtained protein sequences were aligned using Clustal Omega (Sievers et al., 2011) and analyzed for conserved domains. The conserved regions that were targeted by the individual primers as well as corresponding reference entries from the NCBI Conserved Domain Database (Marchler-Bauer et al., 2015) are listed in Supplementary Table S1. Potential oligonucleotide primers were designed using the CODEHOP program (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) (Rose et al., 2003). Sequences of primers developed in this study are listed in Table 1.

2.4. Amplicon sequencing

Next generation amplicon sequencing was performed for seven primer pairs using the Illumina MiSeq platform (Illumina Inc., USA). Library preparation basically followed the protocol "16S Metagenomic Sequencing Library Preparation" (Illumina Inc., USA) but included further adaption steps to the specific characteristics of the primers. Briefly, a first round of amplicon PCR (PCR1) was performed using target specific primers (Table 1) and genomic DNA extracted from forest soil. The reaction was performed in triplicates and comprised: 2.5 μl 10× FastStart High Fidelity Reaction Buffer (Roche Diagnostics, Germany), 2 μl dNTPs (2 nmol each), 1 μl of each primer (10 pmol), 0.5 μl of BSA (3%), 1.25 U FastStart High Fidelity Enzyme Blend (Roche Diagnostics, Germany), 25 ng of genomic DNA (50 ng for *appA* and *phoN*) and ad water 25 μl. The PCR cycling conditions are listed below: Hotstart

Table 1
Listed are investigated genes, classification of corresponding proteins, the mean amplicon length and nucleotide sequences of developed primers.

Protein	Classification	Target gene	Mean expected amplicon length (bp)	Reference	Primer name	Primer sequence 5′–3′
Alkaline phosphatase (phoD)	EC 3.1.3.1	*phoD*	208	Eder et al. (1996)	*phoD*-FW	TGTTCCACCTGGGCGA YWMIATHTAYG
					phoD-RW	CGTTCGCGACCTCGTG RTCRTCCCA
Acid phosphatase (class A)	EC 3.1.3.2	*phoN*	159	Rossolini et al. (1998)	*phoN*-FW	GGAAGAACGGCTCCTA CCCIWSNGGNCA
					phoN-RW	CACGTCGGACTGCCAG TGIDMIYYRCA
6-Phytase/acid phosphatase	EC 3.1.3.26/EC 3.1.3.2	*appA*	375	Golovan et al. (2000)	*appA*-FW	AGAGGGTGGTGATCGTGATG MGICAYGGNRT
					appA-RW	GCCTCGATGGGGTTGA AIADNGGRTC
Phosphonoacetaldehyde hydrolase	EC 3.11.3.1	*phnX*	147	McGrath et al. (2013)	*phnX*-FW	CGTGATCTTCGACtGGGCNGGNAC
					phnX-RW	GTGGTCCCACTTCCCC ADICCCATNGG
Quinoprotein glucose dehydrogenase	EC 1.1.5.2	*gcd*	330	Cleton-Jansen et al. (1990)	*gcd*-FW	CGGCGTCATCCGGGSITIYRAYRT
					gcd-RW	GGGCATGTCCATGTCC CAIADRTCRTG
Phosphate inorganic transporter	TC 2.A.20	*pitA*	270	Elvin et al. (1986)	*pitA*-FW	GGTCTTCGAGTTCATGAACG GNTTYCAYGA
					pitA-RW	CCAGGTGACCAGGTTCCAIRNDAT
Phosphate-specific transporter (periplasmic phosphate-binding protein)	TC 3.A.1.7.1/EC 3.6.3.27	*pstS*	221	Hsieh and Wanner (2010)	*pstS*-FW	TCTACCTGGGGAAGATCACA AARTGGRAYGA
					pstS-RW	TGCCGACGGGCCAITYNWC

(95 °C; 7 min), 30 cycles of denaturation (95 °C; 1 min), annealing (60 °C; 1 min) and elongation (72 °C; 45 s) followed by a final elongation step (72 °C; 7 min). The optimum annealing temperature of the individual primers was investigated in a preliminary test performing a gradient PCR (annealing temperature: 55 °C to 65 °C). Further on a second amplicon PCR (PCR2) was performed using target specific primers comprising Illumina overhang adapter sequences. The amplification was conducted as previously stated. However 2 μl of PCR1 were used as template, 10 pmol of each primer containing adapter overhangs were applied and the cycle number was limited to ten cycles. The amplified triplicates (PCR2) were analyzed on a 2% agarose gel, pooled and subsequently purified using the PCR clean-up /Gel extraction kit (Macherey-Nagel, Germany) according to the manufacturer's protocol. However the washing-buffer (NTI) was diluted in a 1:4 ratio with DEPC treated water prior to elution. Regarding the *phoD* primers three distinct bands were detected after amplification (Supplementary Fig. S1). To check whether all bands represented the target gene the three bands were extracted separately from the agarose gel (PCR clean-up /Gel extraction kit; Macherey-Nagel, Germany) and subsequently treated as individual amplicons (datasets named: *phoD*-K, *phoD*-M, *phoD*-L). Additionally *phoD* amplicons were purified directly from the PCR2 reaction in a second approach (without separating individual bands) (dataset named: *phoD*). Moreover negative controls of the PCR (NKP) as well as negative controls of the DNA-extraction procedure (NKE) were performed for all primers. Finally a total of ten amplicon samples (Supplementary Table S2) plus sixteen negative controls were further processed after PCR2. The correct amplicon sizes were checked on a Bioanalyzer 2100 instrument (Agilent Technologies, USA) using the DNA 7500 kit (Agilent Technologies, USA). The concentration of the purified samples was measured by the Quant-iT PicoGreen kit (Life Technologies, USA). For library preparation the Nextera XT v2 Index kit set A was used (Illumina Inc., USA). The Indexing PCR was performed in 25 μl reactions containing: 12.5 μl NEBNext High-Fidelity Master Mix (New England Biolabs, USA), 2.5 μl of each Indexing primer, 10 ng of purified amplicons and 6.5 μl DECP treated water. The amplification procedure included an initial denaturation step (98 °C; 30 s), 8 cycles of denaturation (98 °C; 10 s), annealing (55 °C; 30 s) and elongation (72 °C; 30 s) followed by a final extension step (72 °C; 5 min). The amplicons were checked on a 2% agarose gel, purified as mentioned previously, analyzed on a Bioanalyzer 2100 instrument (Agilent Technologies, USA) using a DNA 7500 chip (Agilent Technologies, USA) and finally quantified using the Quant-iT PicoGreen kit (Life Technologies, USA). Amplicons were pooled equimolar to 4 nM and sequenced using the MiSeq Reagent kit v3 (600 cycles) (Illumina Inc., USA) for paired end sequencing.

2.5. Amplicon sequencing analysis

FASTQ files of sequenced amplicons were trimmed and merged using AdapterRemoval v2 (minimum read length = 50; minimum Phred quality = 15) (Schubert et al., 2016). Datasets were quality filtered applying the QIIME suite (version 1.9.1) (Caporaso et al., 2010) calling the script "*split_libraries_fastq.py*" (minimum per read length fraction = 0.01; Phred quality threshold = 20; length filtering = 100–600 bp). Details of the sequencing run including number of merged reads and percentage of high quality sequences per dataset are summarized in Supplementary Table S2. All samples were subsampled to 130,377 sequences according to the lowest number of reads obtained within the datasets after quality filtering (Supplementary Table S2) (http://biopieces.org). Subsequently open reading frames (ORF) were predicted utilizing FragGeneScan (version 1.19) (train = illumina_5; thread = 15) (Rho et al., 2010). For functional annotation of datasets hmmsearch (http://hmmer.org; HMMER 3.0) was performed on predicted ORF using the Pfam (version 27.0) (Finn et al., 2014) respectively TIGRFAMs database (version 15) (Haft et al., 2013). Subsequently overlapping profile Hidden Markov Models (HMMs) were removed and results were quality filtered (expect value = 10^{-3}). The quantity of predicted ORF per dataset and the total number of HMMs detected per dataset before and after quality filtering are listed in Table 2 (based on the entire Pfam respectively TIGRFAMs database). To assess the specificity of the primers with respect to the individual targeted genes, the relative percentage of particular HMMs is stated that specifically comprise conserved domains of the proteins encoded by the investigated genes (percentage related to the total number of HMMs detected; Table 2). For taxonomic annotation of the datasets all "positive" sequences (i.e. sequences comprising the conserved domain of the respective HMM)

Table 2
Listed are number of predicted open-reading frames (ORF) per dataset, total number of detected HMMs before and after quality filtering as well as number and percentage of specific HMMs with conserved domains of targeted genes.

Amplicon name	Database	Number of predicted ORF	Total number of HMMs (complete database)	Total number of filtered HMMs (complete database)	Specific HMM	Number of specific HMMs	Percentage of specific HMMs
phoD	Pfam	130,500	125,702	117,210	PF09423	117,161	99.96
phoD-K	Pfam	130,431	121,727	113,543	PF09423	113,504	99.97
phoD-M	Pfam	130,428	122,100	114,914	PF09423	114,888	99.98
phoD-L	Pfam	130,529	133,071	121,738	PF09423	121,463	99.77
phoN	Pfam	129,784	119,958	87,388	PF01569	75,736	86.67
appA	Pfam	131,002	239,090	123,070	PF00328	120,980	98.30
phnX	TIGRFAMs	130,333	256,412	128,376	TIGR01422	128,189	99.85
gcd	TIGRFAMs	131,090	466,310	130,909	TIGR03074	130,536	99.72
pitA	Pfam	130,560	157,192	130,356	PF01384	130,351	100.00
pstS	TIGRFAMs	130,415	248,069	130,264	TIGR00975	130,252	99.99

were aligned against the NCBI RefSeq database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/; February 2016) (Tatusova et al., 2014) using DIAMOND with default settings (Buchfink et al., 2015). Since the *appA* amplicons reached an average length of 375 bp the "sensitive" option in DIAMOND was applied for all datasets. DIAMOND results were imported into MEGAN (version 5.6.5) (Huson et al., 2011) and analyzed on taxonomic level applying the parameters given below: Min Score: 50, Max Expected: 10^{-4}, Top Percent: 10, Min Support Percent: 0, Min Support: 1, LCA Percent: 50, Min complexity: 0.0. Rarefaction analysis of taxonomically annotated datasets was implemented using the R environment (R Core Team, 2015) and package "vegan" (Oksanen et al., 2015). Sequences are stored in the European Nucleotide Archive (Accession number: PRJEB13290).

3. Results and discussion

3.1. Primer specificity

In total seven oligonucleotide primers were developed for genes that code for enzymes which perform major processes of the soil microbial phosphorus (P) turnover. For the entire set of primers a high degree of specificity was determined during amplification of genomic DNA extracted from forest soil. The PCR consistently generated distinct bands of the expected size while unspecific amplification was not observed (Supplementary Fig. S1). Marginal levels of smear around the central band were caused by the amplification procedure since two different types of the individual primers were used in preparation for Illumina sequencing. While the first round of amplicon PCR (PCR1) was performed with the target specific primers (Table 1), the subsequent PCR (PCR2) required Illumina specific adapter overhangs which increased the mean amplicon length by approximately 70 bp. Minor levels of "shorter" fragments from PCR1 were still detectable after PCR2. Regarding the *phoD* gene three distinct amplicon sizes (bands) were detected after amplification (Supplementary Fig. S1). To investigate whether all of them corresponded to the *phoD* gene the bands were separated and subsequently treated as individual amplicons.

The specific amplification of the respective target genes was demonstrated by amplicon sequencing on an Illumina MiSeq instrument. After quality filtering 207,661 sequences were obtained per dataset on average (±37,263) (Supplementary Table S2). Open-reading frames (ORF) were predicted in subsampled datasets (130,377 sequences; according to the lowest number obtained in one dataset) and subsequently scanned against the Pfam (Finn et al., 2014) respectively TIGRFAMs database (Haft et al., 2013) (Table 2). After quality filtering the smallest number of profile Hidden Markov Models (HMMs) was detected in the *phoN* dataset (87,388) while for the other genes 123,375 HMMs on average were obtained (Table 2). Subsequently the relative percentage of specific HMMs, which comprised conserved domains of the proteins encoded by the investigated genes, was determined (Table 2). The sequences that gave a match to the respective (specific) HMM were considered as correctly amplified target genes. After all a satisfactorily high rate of primer specificity was proven. Regarding the primers for *phoD*, *phnX*, *gcd*, *pitA* and *pstS* more than 99% of all sequences that contained an ORF corresponded to the respective target gene (Table 2); for the *appA* primers a specificity of 98% was reached; only for *phoN* specificity levels were below 90% (87%). Regarding the unspecific sequences (*phoN*) the majority of reads represented genes coding for phosphopyruvate hydratases (Pfam family: PF00113). The corresponding proteins comprise certain conserved amino acids within the signature sequence motif of class A acid phosphatases (G-S-Y-P-S-G-H-T) which is targeted by the *phoN*-FW primer (Supplementary Table S1). Since the degree of primer degeneracy is relatively high (*phoN*-FW: 64×) to cover a broad diversity of different microbes, a minor fraction of unspecific sequences was amplified. However the vast majority of amplicons corresponded to the target gene. Therefore the latter primer can also be considered as a suitable tool for the detection of microbes harboring the *phoN* gene. Remarkably the three distinct bands that were amplified by the *phoD* primers (Supplementary Fig. S1) uniformly represented the *phoD* gene. The primer specificity towards the *phoD* gene consistently reached 99% for the individual datasets *phoD*-K, *phoD*-M and *phoD*-L (Table 2). This underlines the broad diversity of microbial *phoD* genes existing in environmental samples which has been previously described (Bergkemper et al., 2015). The amplicon size of *phoD*-K and *phoD*-L differed by sixty base pairs on average. The variable region is likely to be situated in the front section of the *phoD* gene since the primers amplified the upstream region of the *phoD* conserved domain W-D-D-H-E (Supplementary Table S1) (position 621 to 831 of *phoD* from *Bacillus subtilis* subsp. subtilis str. 168; Accession number: gi|255767073). In contrast the PHOD primers introduced by Ragot et al. (2015) target the downstream region of the *phoD* conserved domain; here the occurrence of distinct amplicon sizes was not reported.

Due to the high degree of specificity towards the individual target genes together with the medium amplicon size ranging from 147 bp to 375 bp (Table 1) the introduced primers are not only suitable for diversity analysis but also for quantitative real-time PCR approaches (Karsai et al., 2002).

3.2. Microbial diversity of amplified genes

Insights into the microbial community harboring the investigated genes were gained by taxonomic annotation of amplicon datasets using DIAMOND (Buchfink et al., 2015) against the NCBI RefSeq database (Tatusova et al., 2014) and MEGAN (Huson et al., 2011). Rarefaction analysis based on the number of assigned orders revealed a sufficient coverage of the microbial diversity for all primers (Fig. 1). Indeed the rarefaction curves for *phnX* and *appA* reached a plateau already after 13 detected microbial orders (corresponding to 80,000 reads analyzed), while for *phoD*-L saturation was reached not before 42 detected orders respectively 100,000 analyzed sequences. Thus additional sequencing effort is unlikely to increase the detected microbial diversity in the investigated soil.

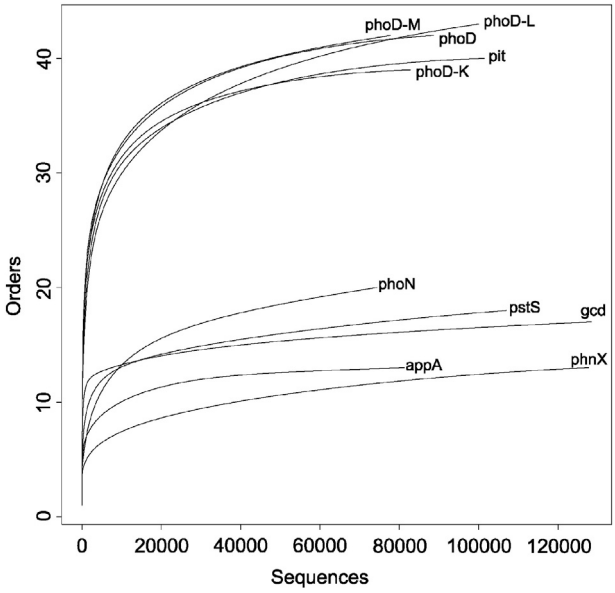

Fig. 1. Rarefaction analysis. Depicted is the number of microbial orders as a function of sequencing depth that were detected in amplicon datasets using seven newly designed oligonucleotide primers and forest soil genomic DNA. Sequences were aligned against the NCBI RefSeq database applying DIAMOND and MEGAN5.

Table 3
Listed are number of detected microbial phyla, orders and families that were amplified by seven oligonucleotide primers from forest soil genomic DNA. Amplified sequences were aligned against the NCBI RefSeq database (DIAMOND) and analyzed using MEGAN5.

Amplicon name	Gene	Number of detected		
		Phyla	Orders	Families
phoD	*phoD*	15	42	77
phoD-K	*phoD*	14	39	69
phoD-M	*phoD*	14	42	78
phoD-L	*phoD*	15	43	81
phoN	*phoN*	7	20	36
appA	*appA*	2	13	13
phnX	*phnX*	2	13	19
gcd	*gcd*	6	17	24
pitA	*pitA*	13	40	68
pstS	*pstS*	9	18	28

In Fig. 2 the relative abundance of microbial orders that were amplified with the individual primers is shown. The highest microbial diversity was detected for the *phoD* and *pitA* genes. After all genes from 15 respectively 13 different microbial phyla were amplified which corresponded to 43 (40) orders or 81 (68) distinct families (Table 3). Dominating orders for *phoD* were Planctomycetales (33%), Burkholderiales (16%), Rhizobiales (14%), Pseudomonadales (12%), Rhodospirillales (9%) and Actinomycetales (6%); 25 orders had a relatively low abundance of 1% or less. In comparison WGS of the investigated soil and filtering of reads related to *phoD* genes revealed a dominance of Acidobacteriales, Rhizobiales, Burkholderiales, Rhodospirillales, Xanthomonadales and Spartobacteria (Bergkemper et al., 2015). Hence the new *phoD* primers covered all taxa that were revealed by WGS (although not necessarily with the same abundance) and further indicated the impact of 38 additional microbial orders on organic-P mineralization by alkaline phosphatase (PhoD) in soil. While the primers were designed based on bacterial sequences exclusively also six fungal orders were detected in the *phoD* dataset: Pleosporales, Chaetothyriales, Onygenales, Hypocreales, Sordariales and Dacrymycetales. In comparison Ragot et al. (2015) amplified the *phoD* gene from 13 different (exclusively) bacterial phyla respectively 38 orders using their primer system. While dominating orders were similar to the presented study, the authors detected also members of Bacillales and Gloeobacterales harboring *phoD* genes in soils. However taking into account that Ragot et al. (2015) sequenced DNA from six different grassland soils with a consistently lower sequencing depth (<5000 sequences per library), the results are not absolutely transferable but may benchmark the efficiency of our newly developed *phoD* primers and validate their efficiency. Regarding the three distinct fragment sizes amplified by the *phoD* primers the underlying microbial community differed for *phoD*-K respectively *phoD*-L (Fig. 2). While longer amplicon sizes (*phoD*-L) predominantly corresponded to a higher abundance of Burkholderiales, Rhizobiales and Pseudomonadales shorter fragments (*phoD*-K) primarily represented genes harbored by Planctomycetales and Rhodospirillales.

Using the newly developed primer system for *phoN* genes 20 different microbial orders were amplified including Rhizobiales (57%), Xanthomonadales (29%), Enterobacteriales (8%) and Burkholderiales (6%), which corresponds again nicely with the WGS dataset from Bergkemper et al. (2015) where the dominance of Rhizobiales and Xanthomonadales in the group of *phoN* harboring bacteria was shown for the same soil. Since sequencing depth has been limited for WGS again additional orders were detected by the amplicon based approach, which were not described in the metagenomes.

Microbial enzymes that degrade soil organic myo-Inositol-1,2,3,4,5,6-hexakisphosphates (IP_6) have been classified as 3-phytases

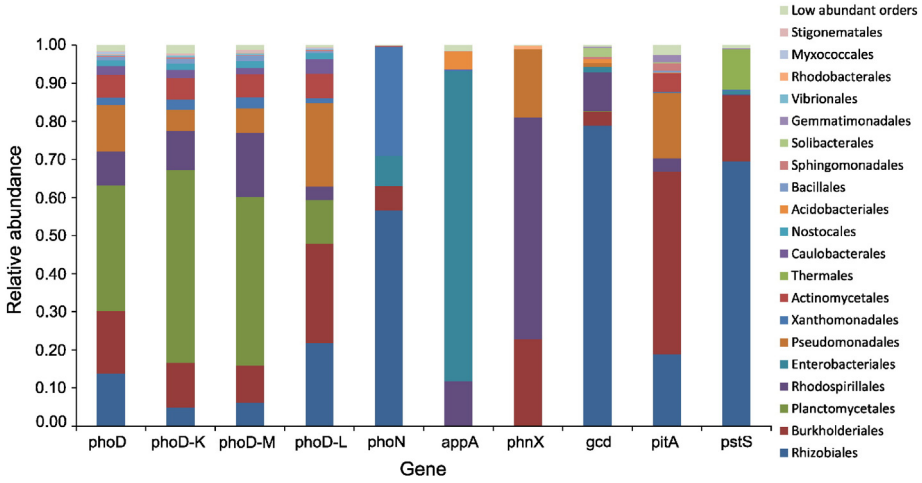

Fig. 2. Taxonomic diversity of amplified genes. Seven newly designed oligonucleotide primers were applied for amplicon sequencing of forest soil genomic DNA. Sequences of amplified target genes were aligned against the NCBI RefSeq database using DIAMOND and analyzed on taxonomic level using MEGAN.

and 6-phytases depending on the initial dephosphorylation (Jorquera et al., 2008). The primers introduced in this study specifically targeted the *appA* gene, which codes for a bifunctional enzyme with phytase and acid phosphatase activity that is characterized as histidine-acid phosphatase (HAP; 6-phytase) (Golovan et al., 2000). In total the *appA* primers amplified genes from 2 different bacterial phyla and 13 distinct orders (Table 3). Compared to the microbial diversity that was covered by the *phoD* primers this number appeared to be insufficient given the fact that IP_6 is considered as the major storage form of soil organic P (Turner et al., 2002). However HAP represents only one single class of phytases beyond beta propeller phytases (BPP) and purple acid phosphatases (PAP). Unfortunately conserved domains suitable for primer design were only present in HAP (Mullaney and Ullah, 2003). Moreover database entries of sequenced HAP are rare and generally restricted to members of Alpha-, Delta-, Gammaproteobacteria and Acidobacteria (Jorquera et al., 2008). Therefrom the *appA* primers amplified genes from all classes of microorganisms that are known for harboring HAP genes (Fig. 2).

Similarly a relatively low microbial diversity was amplified by the *phnX* primers, which target the gene coding for a phosphonatase (McGrath et al., 2013). Two different bacterial phyla (Proteobacteria, Firmicutes) and 13 orders were detected. Database entries of *phnX* genes are limited (<10% compared to *phoD*; NCBI RefSeq; March 2016) and comprise predominantly Actinobacteria, CFB group bacteria (Cytophaga–Flavobacter–Bacteroides), Proteobacteria and Firmicutes. The latter two phyla were covered by the *phnX* primers while Actinobacterial sequences were not amplified (Fig. 2). This is caused by the fact that phosphonatases are not well conserved on amino acid level when comparing microorganisms from distinct phyla. Therefore the *phnX* primers were exclusively designed based on sequences derived from different classes of Proteobacteria and Firmicutes while Actinobacteria and CFB group bacteria were omitted. In a preliminary experiment, an additional set of highly degenerated primers was investigated that also targeted the *phnX* gene of Actinobacteria. As degeneracy increased manifold the PCR failed to produce target specific amplicons (data not shown). Therefore *phnX* primers were designed to reliably amplify target genes while the primer degeneracy was minimized.

In contrast to the enzymes previously discussed which are involved in soil organic-P mineralization the quinoprotein glucose dehydrogenase (PQQGDH; *gcd*) facilitates the solubilization of mineral-P (Goldstein, 1994) and serves as an indicator for the microbial potential of inorganic-P solubilization. The *gcd* primers amplified target genes from 6 distinct bacterial phyla and 17 orders. Most abundant were genes harbored by Rhizobiales (79%), Rhodospirillales (10%) and Burkholderiales (4%) which are known for their ability to solubilize inorganic-P compounds (Rodriguez and Fraga, 1999) (Fig. 2). The strikingly high abundance of Rhizobial *gcd* genes certainly is influenced by the strong domination of this order in the investigated soil (Bergkemper et al., 2015). WGS revealed also members of Solibacterales and Acidobacteriales as important sources for *gcd* in BBR. Both orders were likewise covered by the *gcd* primers although their relative abundance (2% respectively 1% of amplified sequences) was low compared to the WGS approach (Bergkemper et al., 2015). On the one hand this could be related to an overrepresentation of Proteobacterial *gcd* sequences in common databases (>80% of *gcd* sequences in NCBI RefSeq; March 2016). On the other hand microorganisms are highly stratified in the different forest soil horizons (Baldrian et al., 2012). Hence the *gcd* harboring microbial community might shift from the soil organic horizon used in the present study to the mineral topsoil applied for WGS (Bergkemper et al., 2015) especially since inorganic-P solubilization might play a major role once the content of soil organic-P is decreased (e.g. in mineral soil horizons) (Talkner et al., 2009).

Genes coding for the phosphate inorganic transporter (Pit) were amplified from a broad diversity of microorganisms including 13 different bacterial phyla and 40 orders (Table 3). Dominating taxa were Burkholderiales (48%), Rhizobiales (19%), Pseudomonadales (17%), Actinomycetales (5%) and Rhodospirillales (3%) (Fig. 2). In addition all remaining orders that were known for harboring the *pitA* gene in the investigated soil were covered by the *pitA* primers (Acidobacteriales, Solibacterales, Myxococcales, Gemmatimonadales) (Bergkemper et al., 2015). The gene *pstS* that codes for the periplasmic binding protein of the phosphate-specific transporter (Pst) was amplified from 9 different bacterial phyla respectively 18 orders (Table 3). Most abundant were members of Rhizobiales (70%), Burkholderiales (17%), Thermales (11%) and Xanthomonadales (0.4%) (Fig. 2). Therefrom the primers covered 7 (of 11) *pstS* harboring orders that were detected by WGS while the target gene was amplified from 11 additional orders using the *pstS* primers.

4. Conclusion

In conclusion a new set of oligonucleotide primers was introduced in this study that covers the major processes of the soil microbial phosphorus turnover. The seven primers target genes which code for proteins involved in mineralization of various forms of soil organic-P, solubilization of inorganic-P as well as cellular P uptake. A novel strategy for primer design was applied to allow both the amplification of target genes from a broad diversity of distinct microorganisms and simultaneously minimize the grade of primer degeneracy. Microbial key players that strongly contributed to the turnover of soil P in beech forests were identified by whole genome shotgun sequencing while corresponding database sequences of the individual enzymes were used for primer design. The new primers showed a high degree of specificity during amplification while simultaneously a broad diversity of distinct microbial phyla was covered and the results obtained by metagenomics were nearly reproduced. On this account the introduced primers represent a valuable tool to deepen our knowledge on key players of this crucial nutrient turnover in soils and other environments. Moreover the primers are suitable for application in quantitative real-time PCR approaches which allows the quantification of important transformation steps in P turnover both on temporal and spatial scale using DNA (as a proxy for potentials) and mRNA (as a proxy for activity).

Acknowledgments

Fabian Bergkemper was supported by the German Research Foundation (DFG) (SCHL 446/20-1) in frame of the Priority Program “Ecosystem Nutrition: Forest Strategies for limited Phosphorus Resources” (SPP 1685). The project was further supported by the BMBF funded project “InnoSoilPhos”.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.mimet.2016.04.011.

References

Baldrian, P., Kolarik, M., Stursova, M., Kopecky, J., Valaskova, V., Vetrovsky, T., Zifcakova, L., Snajdr, J., Ridl, J., Vlcek, C., Voriskova, J., 2012. Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. ISME J. 6, 248–258.

Bergkemper, F., Schöler, A., Engel, M., Lang, F., Krüger, J., Schloter, M., Schulz, S., 2015. Phosphorus depletion in forest soils shapes bacterial communities towards phosphorus recycling systems. Environ. Microbiol.

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60.

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

Cleton-Jansen, A.M., Goosen, N., Fayet, O., van de Putte, P., 1990. Cloning, mapping, and sequencing of the gene encoding *Escherichia coli* quinoprotein glucose dehydrogenase. J. Bacteriol. 172, 6308–6315.
R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Delmont, T.O., Robe, P., Cecillon, S., Clark, I.M., Constancias, F., Simonet, P., Hirsch, P.R., Vogel, T.M., 2011. Accessing the soil metagenome for studies of microbial diversity. Appl. Environ. Microbiol. 77, 1315–1324.
Eder, S., Shi, L., Jensen, K., Yamane, K., Hulett, F.M., 1996. A *Bacillus subtilis* secreted phosphodiesterase/alkaline phosphatase is the product of a Pho regulon gene, *phoD*. Microbiology 142, 2041–2047.
Elser, J.J., 2012. Phosphorus: a limiting nutrient for humanity? Curr. Opin. Biotechnol. 23, 833–838.
Elvin, C., Dixon, N., Rosenberg, H., 1986. Molecular cloning of the phosphate (inorganic) transport (pit) gene of *Escherichia coli* K12. Mol. Gen. Genet. 204, 477–484.
Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L., Tate, J., Punta, M., 2014. Pfam: the protein families database. Nucleic Acids Res. 42, D222–D230.
Goldstein, A., 1994. Involvement of the quinoprotein glucose dehydrohenase in the solubilization of exogenous phosphates by Gram-negative bacteria. Phosphate in microorganisms: Cellular and molecular biology, pp. 197–203.
Golovan, S., Wang, G., Zhang, J., Forsberg, C.W., 2000. Characterization and overproduction of the *Escherichia coli appA* encoded bifunctional enzyme that exhibits both phytase and acid phosphatase activities. Can. J. Microbiol. 46, 59–71.
Haft, D.H., Selengut, J.D., Richter, R.A., Harkins, D., Basu, M.K., Beck, E., 2013. TIGRFAMs and Genome Properties in 2013. Nucleic Acids Res. 41, D387–D395.
Holford, I.C.R., 1997. Soil phosphorus: its measurement, and its uptake by plants. Soil Res. 35, 227–240.
Hsieh, Y.-J., Wanner, B.L., 2010. Global regulation by the seven-component Pi signaling system. Curr. Opin. Microbiol. 13, 198–203.
Huson, D.H., Mitra, S., Ruscheweyh, H.J., Weber, N., Schuster, S.C., 2011. Integrative analysis of environmental sequences using MEGAN4. Genome Res. 21, 1552–1560.
Jorquera, M., Martinez, O., Maruyama, F., Marschner, P., de la Luz Mora, M., 2008. Current and future biotechnological applications of bacterial phytases and phytase-producing bacteria. Microbes Environ./JSME. 23, 182–191.
Karsai, A., Müller, S., Platz, S., Hauser, M.-T., 2002. Evaluation of a Homemade SYBR(®) Green I Reaction Mixture for Real-Time PCR Quantification of Gene Expression. Biotechniques 32, 790–796.
Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Bryant, S.H., 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43, D222–D226.
McGrath, J.W., Chin, J.P., Quinn, J.P., 2013. Organophosphonates revealed: new insights into the microbial metabolism of ancient molecules. Nat. Rev. Microbiol. 11, 412–419.
Mullaney, E.J., Ullah, A.H.J., 2003. The term phytase comprises several different classes of enzymes. Biochem. Biophys. Res. Commun. 312, 179–184.
Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2015. vegan: Community Ecology Package. R package version 2.3-1.
Ragot, S.A., Kertesz, M.A., Bunemann, E.K., 2015. *phoD* Alkaline Phosphatase Gene Diversity in Soil. Appl. Environ. Microbiol. 81, 7281–7289.
Rho, M., Tang, H., Ye, Y., 2010. FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acids Res. 38, e191.
Richardson, A.E., Simpson, R.J., 2011. Soil microorganisms mediating phosphorus availability. Plant Physiol. 156, 989–996.
Rodriguez, H., Fraga, R., 1999. Phosphate solubilizing bacteria and their role in plant growth promotion. Biotechnol. Adv. 17, 319–339.
Rose, T.M., Henikoff, J.G., Henikoff, S., 2003. CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. Nucleic Acids Res. 31, 3763–3766.
Rossolini, G.M., Schippa, S., Riccio, M.L., Berlutti, F., Macaskie, L.E., Thaller, M.C., 1998. Bacterial nonspecific acid phosphohydrolases: physiology, evolution and use as tools in microbial biotechnology. CMLS, Cell. Mol. Life Sci. 54, 833–850.
Sakurai, M., Wasaki, J., Tomizawa, Y., Shinano, T., Osaki, M., 2008. Analysis of bacterial communities on alkaline phosphatase genes in soil supplied with organic matter. Soil Sci. Plant Nutr. 54, 62–71.
Schöler, A., de Vries, M., Vestergaard, G., Schloter, M., 2016. Reconstruction of Transformation Processes Catalyzed by the Soil Microbiome Using Metagenomic Approaches. Microbial Environmental Genomics (MEG)Springer, New York, pp. 197–206.
Schubert, M., Lindgreen, S., Orlando, L., 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Research Notes. 9, p. 88.
Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7.
Talkner, U., Jansen, M., Beese, F.O., 2009. Soil phosphorus status and turnover in central-European beech forest ecosystems with differing tree species diversity. Eur. J. Soil Sci. 60, 338–346.
Tan, H., Barret, M., Mooij, M., Rice, O., Morrissey, J., Dobson, A., Griffiths, B., O'Gara, F., 2013. Long-term phosphorus fertilisation increased the diversity of the total bacterial community and the *phoD* phosphorus mineraliser group in pasture soils. Biol. Fertil. Soils 49, 661–672.
Tatusova, T., Ciufo, S., Fedorov, B., O'Neill, K., Tolstoy, I., 2014. RefSeq microbial genomes database: new representation and annotation strategy. Nucleic Acids Res. 42, D553–D559.
Töwe, S., Wallisch, S., Bannert, A., Fischer, D., Hai, B., Haesler, F., Kleineidam, K., Schloter, M., 2011. Improved protocol for the simultaneous extraction and column-based separation of DNA and RNA from different soils. J. Microbiol. Methods 84, 406–412.
Turner, B.L., Papházy, M.J., Haygarth, P.M., McKelvie, I.D., 2002. Inositol phosphates in the environment. Philos. Trans. R. Soc. B 357, 449–469.
Vitousek, P.M., Porder, S., Houlton, B.Z., Chadwick, O.A., 2010. Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen-phosphorus interactions. Ecological Applications: a Publication of the Ecological Society of America. 20, pp. 5–15.