



# Functional Metagenomics Reveals an Overlooked Diversity and Novel Features of Soil-Derived Bacterial Phosphatases and Phytases

Genis Andrés Castillo Villamizar,<sup>a,b</sup> Heiko Nacke,<sup>a</sup> Marc Boehning,<sup>a\*</sup> Kristin Herz,<sup>a</sup>  Rolf Daniel<sup>a</sup>

<sup>a</sup>Department of Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University, Göttingen, Germany

<sup>b</sup>Línea tecnológica biocorrosión, Corporación para la investigación de la corrosión C.I.C. Piedecuesta, Santander, Colombia

**ABSTRACT** Phosphatases, including phytases, play a major role in cell metabolism, phosphorus cycle, biotechnology, and pathogenic processes. Nevertheless, their discovery by functional metagenomics is challenging. Here, soil metagenomic libraries were successfully screened for genes encoding phosphatase activity. In this context, we report the largest number and diversity of phosphatase genes derived from functional metagenome analysis. Two of the detected gene products carry domains which have never been associated with phosphatase activity before. One of these domains, the SNARE-associated domain DedA, harbors a so-far-overlooked motif present in numerous bacterial SNARE-associated proteins. Our analysis revealed a previously unreported phytase activity of the alkaline phosphatase and sulfatase superfamily (cl23718) and of purple acid phosphatases from nonvegetal origin. This suggests that the classical concept comprising four classes of phytases should be modified and indicates high performance of our screening method for retrieving novel types of phosphatases/phytases hidden in metagenomes of complex environments.

**IMPORTANCE** Phosphorus (P) is a key element involved in numerous cellular processes and essential to meet global food demand. Phosphatases play a major role in cell metabolism and contribute to control the release of P from phosphorylated organic compounds, including phytate. Apart from the relationship with pathogenesis and the enormous economic relevance, phosphatases/phytases are also important for reduction of phosphorus pollution. Almost all known functional phosphatases/phytases are derived from cultured individual microorganisms. We demonstrate here for the first time the potential of functional metagenomics to exploit the phosphatase/phytase pools hidden in environmental soil samples. The recovered diversity of phosphatases/phytases comprises new types and proteins exhibiting largely unknown characteristics, demonstrating the potential of the screening method for retrieving novel target enzymes. The insights gained into the unknown diversity of genes involved in the P cycle highlight the power of function-based metagenomic screening strategies to study Earth's phosphatase pools.

**KEYWORDS** SNARE-associated domain, functional metagenomics, phosphatases, phytases, soil metagenome

Within the last decades, advances in next-generation sequencing and metagenomic techniques have led to the discovery of new enzymes from metagenomes (1, 2). Novel lipases, esterases, proteases, and hydrogenases, among many others, have been identified (3, 4). Nevertheless, the majority of enzymes with high biological relevance are still almost exclusively recovered from cultured organisms (2).

**Citation** Castillo Villamizar GA, Nacke H, Boehning M, Herz K, Daniel R. 2019. Functional metagenomics reveals an overlooked diversity and novel features of soil-derived bacterial phosphatases and phytases. *mBio* 10:e01966-18. <https://doi.org/10.1128/mBio.01966-18>.

**Editor** Daniel Barkan, Hebrew University of Jerusalem

**Copyright** © 2019 Castillo Villamizar et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Rolf Daniel, [rdaniel@gwdg.de](mailto:rdaniel@gwdg.de).

\* Present address: Marc Boehning, Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany.

**Received** 6 September 2018

**Accepted** 11 December 2018

**Published** 29 January 2019

This is especially the case for phosphatases. Phosphatases have evolved across all living organisms and contribute to the regulation of diverse cellular functions (5, 6). A specific group of phosphatases named phytases can release phosphorus from phytic acid, which is one of the most important phosphorus reserves in plants and soils (7, 8).

Phosphorus (P) reserves are globally important, due to the enormous growth of the world population, and the ensuing demand for this macroelement. Large amounts of P are and will be required in order to fulfill the increasing world agroalimentary needs (9). However, global rock phosphorus reservoirs are currently being rapidly depleted, and the supplementation of P to animal feed and plant fertilizers has become more expensive during the last decades (10). Plant-based animal feeds often contain large amounts of phytate, which cannot be utilized by monogastric animals due to the lack of phytases (7, 11). As a consequence, P levels in soils and water bodies increase. This eutrophication causes for instance algal blooms in aquatic ecosystems, leading to deoxygenated areas disturbing the life of many species (12). To meet future requirements, minimize losses of P, and reduce the environmental impact, it is necessary to use P compounds more efficiently and develop economical recycling technologies. In this context, phosphatases/phytases have proved to be remarkably useful (13). These enzymes are currently used in agroindustry to minimize P losses and to improve the levels of bioavailable P (14). A more recently described role of the phytases is their involvement in pathogenicity causing tissue damage in humans, coordination of the virulence program in *Dickeya dadantii*, and mediation of plant infection by *Candida albicans* and *Xanthomonas*, respectively (5, 15, 16).

The diversity and potential of environmental phytases remain largely unexplored as so far almost all reported functionally characterized phytases were derived from cultured organisms, including plants, fungi, and bacteria. Based on their catalytic characteristics, four classes of phytases have been described: histidine acid phytase (HAPhy),  $\beta$ -propeller phytase (BPPhy), purple acid phytase (PAPhy), and protein tyrosine phytase (PTPhy). These enzymes are structurally and catalytically dissimilar (14, 17).

In this study, we use a function-based screening approach (18) to identify environmental phosphatases/phytases. By using soil metagenomes as a source, we were able to recover novel genes encoding phosphatases with phytase activity. Some of the recovered genes encode protein domains that were not associated with phosphatase activity before, and others represent new types or subtypes of phytases.

## RESULTS

**Phosphatase detection strategy.** The metagenomic libraries contained approximately 38,122 to 166,040 clones and were screened for candidates exhibiting phosphatase activity using plates with phytate as phosphorus source and BCIP as indicator (see Fig. S1 in the supplemental material). The quality of the libraries was controlled by determining the average insert sizes and the percentage of insert-bearing *Escherichia coli* clones. The average insert sizes of metagenomic DNA-containing plasmids ranged from 2.8 to 6.7 kb, and the frequency of clones carrying plasmid inserts was at least 89% (Table 1).

We recovered 21 positive *E. coli* clones from functional screens carrying plasmids harboring one or more ORFs associated with known phosphatase genes and domains (designation of plasmids is given in Table 1). The entire inserts of the positive clones were sequenced and taxonomically classified, showing that in all cases the cloned environmental DNA is of bacterial origin. Most inserts of the positive clones were affiliated with *Terrabacteria*, *Proteobacteria*, and the PVC superphylum with seven, six, and four representatives, respectively. Within the *Terrabacteria* group, most of the inserts (4) were affiliated with *Actinobacteria* (Table S1).

Thirty-one ORFs encoding putative gene products with similarity to known phosphatase enzymes were identified. Signal peptides were detected for 12 of them. The deduced gene products comprised 214 to 819 amino acids with calculated molecular masses ranging from 12 to 65.5 kDa and amino acid sequence identities to the closest

**TABLE 1** Characteristics of the soil metagenomic libraries and designation of plasmids harbored by positive clones

Library <sup>a</sup>	No. of clones	Avg insert size (kb)	Insert frequency (%)	Estimated library size (Gb)	No. of positive clones/Gb	Plasmid(s) of positive clones (accession no.)
AEW1*	129,748	6.7	91	0.79	1.2	pLP01 (KY931670)
AEW5*	90,300	5.2	89	0.42	2.3	pLP02 (KY931671)
SEW2*	135,240	5.7	95	0.73	9.6	pLP10 (KY931677), pLP14 to pLP19 (KY931679 to KY931684)
SEW5*	166,040	4.0	95	0.63	1.6	pLP07 (KY931674)
SEW46	38,122	2.8	93	0.17	23.5	pLP03 (KY931672), pLP04 (KY931673), pLP08 (KY931675), pLP09 (KY931676)
HEW30	53,460	6.1	96	0.31	22.6	pLP13 (KY931678), pLP20 (KY931685), pLP24 to pLP28 (KY931686 to KY931690)

<sup>a</sup>AEW, metagenomic libraries derived from the Biodiversity Exploratory Schwäbische Alb; SEW, metagenomic libraries derived from the Biodiversity Exploratory Schorfheide-Chorin; HEW, metagenomic libraries derived from the Biodiversity Exploratory Hainich-Dün. \*, previously generated libraries (39).

known phosphatases ranging from 25% (Pho14B) to 83% (Pho13) over the full-length protein (Table 2).

From the 21 positive clones, seven harbored more than one putative phosphatase-related gene (Table 2). Thus, if two or more potential phosphatase activity-related genes were present in a positive clone, individual heterologous expression and subsequent phosphatase activity verification were performed. The analysis of colonies showed that the individual heterologous expression of 24 out of 31 genes led to phosphatase activity and the corresponding positive phenotype of the respective recombinant *E. coli* strains (Table 2).

**High phosphatase diversity recovered from soil metagenomes.** Phosphatases can be classified according to the structural fold of the catalytic domains and subclassified into families and subfamilies based on sequence similarities of the phosphatase domains, as well as by conserved amino acid motifs not belonging to the catalytic domain (6, 19). However, some are still classified based on their biochemical properties and biological functions (20).

Among the putative gene products encoded by the 31 candidate genes, alkaline phosphatases were identified as the most abundant group (five representatives), followed by histidine phosphatases and phospholipases with four representatives each. Phosphoserine-phosphatases and protein-tyrosine phosphatases were represented by three putative genes each. Acid phosphatases were encoded by two genes, while the plasmid pLP10 harbored an ORF with a deduced gene product showing similarity to a mismatch repair ATPase (Table 2).

The amino acid sequence analysis revealed the presence of 10 different domains in the 31 deduced proteins. We detected the alkaline phosphatase and sulfatase superfamily domain (ALP-like cl23718) as the most frequent domain, represented in eight sequences. The second highest abundance showed the haloacid dehydrogenase domain (HAD cl21460), which was identified in six protein sequences. Three out of four classical phosphatase/phytase domains were detected in this study: the histidine phosphatase domain (HP with five protein sequences), the tyrosine phosphatase domain (PTPc with two protein sequences), and the acid phosphatase domain (PAP with two protein sequences) (Fig. 1). The phylogenetic analyses of the enzyme sequences and those harboring the above-mentioned domains revealed different clustering patterns in relation to reference phosphatase sequences for the different groups. Within the analyzed groups, the clustering of the metagenome-derived enzymes ranged from clear separation to integrated clustering (Fig. S2).

The HP superfamily (cl11399) is represented by a diverse group of proteins divided into two branches exhibiting numerous functions (21). Classical members of the HAPhy share a conserved motif, RHGXRX, characteristic for this enzyme class. The HAPhy catalytic reactions are based on the conserved histidine residue in the RHGXRX motif (21, 22). In this study, all five phosphatases belonging to the HP superfamily harbored this histidine residue (Fig. 2A). Three out of five HPs in this survey were encoded by

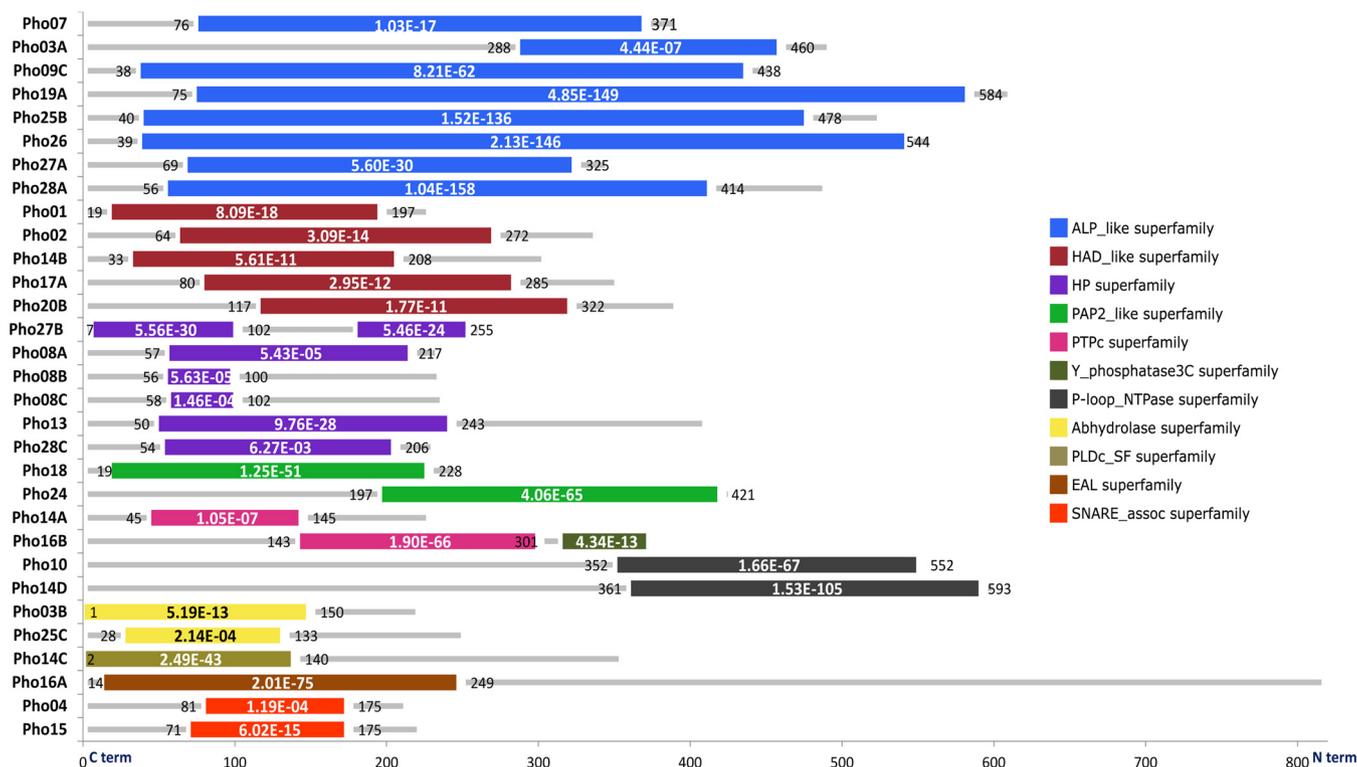
**TABLE 2** Gene products encoded by genes associated with phosphatase activity and their observed sequence identities

Gene (accession no. of protein)	No. of amino acids encoded	Closest similar phosphatase protein, accession no. (no. of encoded amino acids), organism, E value	Identity to closest similar phosphatase protein (Blast), no. of amino acids similar/total no. (%)	% identity to closest similar phosphatase protein (Clustal alignment)
<i>pho01</i> (AWN00218)	229	Phosphatidylglycerophosphatase, PIF15492.1 (224), <i>Rhodanobacter</i> sp. strain TND4EH1, 3E−99	161/213 (76)	72
<i>pho02</i> (AWN00219) <sup>a</sup>	339	Phosphoserine phosphatase, AFM25187 (342), <i>Desulfomonile tiedjei</i> DSM 6799, 0.0	251/337 (74)	74
<i>pho03A</i> (AWN00220) <sup>a</sup>	493	Phosphoesterase, WP_009239878.1 (404), <i>Ralstonia</i> , 2E−9	183/425 (49)	47
<i>pho03B</i> (AWN00221) <sup>b</sup>	222	Phospholipase/carboxylesterase, ADV48687.1 (334), <i>Cellulophaga algicola</i> DSM 14237, 2E−14	84/181 (46)	27
<i>pho04</i> (AWN00222)	214	Putative membrane-associated alkaline phosphatase, KGB26473 (203), <i>Acetobacter tropicalis</i> , 9E−50	92/193 (48)	46
<i>pho07</i> (AWN00223) <sup>a</sup>	392	Phosphoesterase family protein, PZS03611.1 (379), <i>Pseudonocardiales</i> bacterium, 1E−111	184/349 (53)	51
<i>pho08A</i> (AWN00224) <sup>a</sup>	235	Histidine phosphatase family protein, WP_074262886.1 (229), <i>Paraburkholderia phenazinium</i> , 2E−56	97/191 (51)	49
<i>pho08B</i> (AWN00225) <sup>a</sup>	236	Histidine phosphatase family protein, WP_090546752.1 (196), <i>Paraburkholderia caballeronis</i> , 1E−59	97/182 (53)	50
<i>pho08C</i> (AWN00226) <sup>a</sup>	238	Histidine phosphatase family protein, WP_090546752.1 (196), <i>Paraburkholderia caballeronis</i> , 2E−57	98/182 (54)	51
<i>pho09C</i> (AWN00227) <sup>a</sup>	455	Alkaline phosphatase family protein, WP_007415052.1 (407), <i>Pedospaera parvula</i> , 0.0	330/413 (66)	63
<i>pho10</i> (AWN00228)	554	Mismatch repair ATPase, WP_014786775 (599), <i>Terriglobus roseus</i> , 6E−142	246/558 (44)	44
<i>pho13</i> (AWN00229)	411	Broad-specificity phosphatase PhoEn, WP_071949433.1 (401), <i>Mycobacterium</i> sp. strain PYR15, 0.0	349/400 (87)	83
<i>pho14A</i> (AWN00230) <sup>a,b</sup>	229	Protein tyrosine phosphatase (partial), CCZ50566.1 (64), <i>Acidobacteria</i> bacterium, 9E−13	43/111 (50)	48
<i>pho14B</i> (AWN00231) <sup>b</sup>	305	Phosphoserine phosphatase, PKM89459.1 (276), <i>Firmicutes</i> bacterium, 2E−4	58/215 (27)	25
<i>pho14C</i> (AWN00232) <sup>b</sup>	356	Phosphatidylserine/phosphatidyl glycerophosphate, AEQ20292 (371), uncultured bacterium CSLG7, 2E−109	175/357 (49)	48
<i>pho14D</i> (AWN00233) <sup>b</sup>	602	Protein tyrosine phosphatase, PYO70860.1 (581), <i>Gemmatimonadetes</i> bacterium, 1E−137	244/579 (42)	41
<i>pho15</i> (AWN00234)	223	Alkaline phosphatase, OFV86354.1 (209), <i>Acidobacteria</i> bacterium, 8E−34	71/167 (43)	41
<i>pho16A</i> (AWN00235)	819	Diguanylate cyclase/phosphodiesterase, WP_067501625.1 (816), <i>Actinoplanes</i> sp. strain TFC3, 1E−46	105/247 (43)	39
<i>pho16B</i> (AWN00236) <sup>a</sup>	376	Protein tyrosine phosphatase, WP_042381880.1 (372), <i>Streptacidiphilus melanogenes</i> , 0.0	257/324 (89)	76
<i>pho17A</i> (AWN00237) <sup>a</sup>	353	Phosphoserine phosphatase, AFM25187 (342), <i>Desulfomonile tiedjei</i> DSM 6799, 0.0	252/329 (77)	74
<i>pho18</i> (AWN00238) <sup>a</sup>	248	Phosphatase PAP2 family protein, WP_093286091.1 (257), <i>Verrucomicrobiaceae</i> bacterium GAS474, 4E−55	99/200 (50)	46
<i>pho19A</i> (AWN00239)	612	Alkaline phosphatase precursor, AMY11511 (577), <i>Acidobacteria</i> bacterium DSM 100886, 8E−126	230/529 (43)	42
<i>pho20B</i> (AWN00240)	392	Phosphoglycolate phosphatase, RDI59778.1 (337), <i>Microvirga subterranea</i> , 3E−152	248/339 (76)	73
<i>pho24</i> (AWN00241)	428	PAP2 superfamily protein, SHK15444 (414), <i>Bradyrhizobium lablabi</i> , 3E−141	215/405 (53)	54
<i>pho25B</i> (AWN00242) <sup>a,b</sup>	526	Phospholipase, WP_052891151 (505), <i>Thermogemmatispora carboxidivorans</i> , 0.0	303/527 (57)	60
<i>pho25C</i> (AWN00243)	252	Phospholipase, WP_006679394.1 (222), <i>Paenibacillus dendritiformis</i> , 0.0	41/101 (41)	28
<i>pho26</i> (AWN00244) <sup>a</sup>	559	Alkaline phosphatase family protein, WP_020714678.1 (564), <i>Acidobacteriaceae</i> bacterium KBS 89, 0.0	434/551 (79)	78
<i>pho27A</i> (AWN00245) <sup>a</sup>	347	Multispecies: phosphatase, WP_PYV87257.1 (338), <i>Acidobacteria</i> bacterium, 9E−64	249/323 (77)	74
<i>pho27B</i> (AWN00246)	263	Acid sugar phosphatase, GBD30013.1 (265), bacterium HR32, 2E−57	106/254 (42)	39
<i>pho28A</i> (AWN00247) <sup>b</sup>	490	Nonhemolytic phospholipase C, APW61637.1 (486), <i>Paludisphaera borealis</i> , 0.0	328/454 (72)	69
<i>pho28C</i> (AWN00248) <sup>a</sup>	232	Histidine phosphatase family protein, WP_106819986.1 (214), <i>Syntrophobacter</i> sp. strain SbD1, 2E−61	93/170 (53)	46

<sup>a</sup>Signal peptide detected.<sup>b</sup>No phosphatase activity was detected on indicator plates after cloning ORF into expression vector.

plasmid pLP08. The analysis of the plasmid sequence revealed a tandem organization of these genes with slight individual sequence variations (Fig. 2A; Fig. S3).

PTPs are well-studied proteins with a characteristic motif (HCX5R) (23, 24). In this study, two new PTPs (Pho14A and Pho16B) harboring the typical catalytic signature of the group (Fig. 2B) were detected. Interestingly, Pho16B showed the specific signature



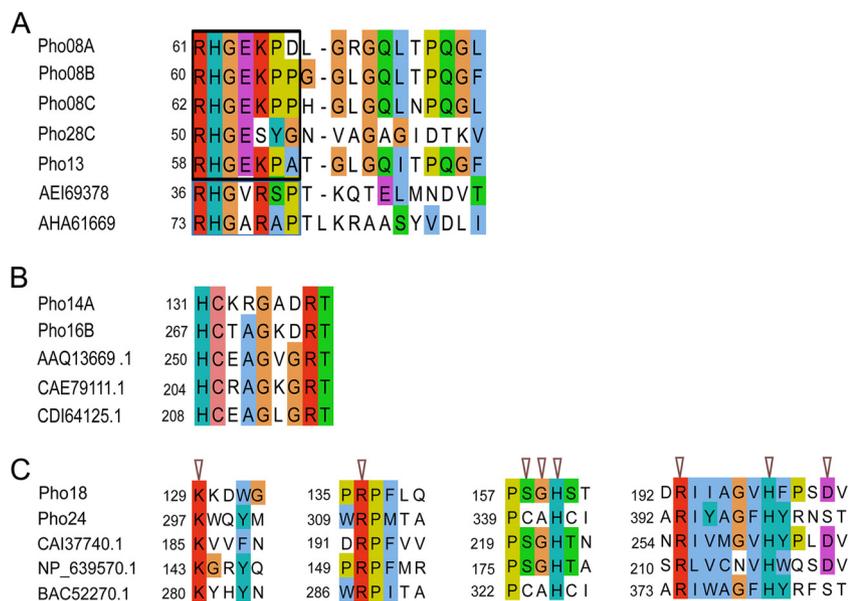
**FIG 1** General architecture and domains of the retrieved phosphatases: ALP, alkaline phosphatases and sulfatases (cl23718); HAD, haloacid dehalogenase (cl21460); HP, histidine phosphatase (cl11399); PAP2, phosphatidic acid phosphatase (cl00474); PTPs, protein tyrosine phosphatases (cl21483); Y phosphatase 3C superfamily (cl6249); P-loop NTPase superfamily (cl21455); abhydrolase superfamily (cl21494); PLDc, phospholipase D (cl15239); EAL superfamily (cl00290); SNARE-associated superfamily (cl00429).

of the MtpB-like phosphatases characterized by the presence of the unique active site P-loop submotif HCXXXKDRT. This type of protein has been predicted in several microorganisms, including pathogens, but never in environmental samples. For the remaining group of classic phytases detected in this study (PAP), the literature describes two branches, the PAP1 enzymes, which are Mg<sup>2+</sup>-dependent enzymes, and the PAP2 enzymes, which are Mg<sup>2+</sup> independent, but in all cases the active forms of PAP phytases were derived from plants (25). We detected the PAPs Pho18 and Pho24, which are affiliated with bacteria and belong to the Mg<sup>2+</sup>-independent branch (PAP2 cl00474) (Fig. 1 and 2C).

Alpha/beta hydrolases (abhydrolases) represent a group of proteins with a high number of substrates and catalytic functions (26). Two gene products (Pho03B and Pho25C) contained an abhydrolase domain (Fig. 1). However, only Pho25C showed phosphatase/phytase activity after individual heterologous expression of the corresponding gene. Abhydrolases exhibit broad substrate specificity, and some members have been reported with phospholipase activity (27).

Other ORFs such as Pho16A carry the EAL domain, which is present in diverse bacterial signaling proteins and encodes a phosphodiesterase function (28). Analysis of Pho10 and Pho14D amino acid sequences indicates the presence of the P-loop\_NTPase superfamily domain (Fig. 1). Enzymes harboring this domain hydrolyze the beta-gamma phosphate bond of, e.g., ATP and GTP (29). In this study, Pho10 showed phosphatase activity, while Pho14D as part of the clone harboring plasmid pLP14 showed none. Pho14C showed no phosphatase activity after individual heterologous expression of the corresponding gene. The *pho14C* gene product harbors the phospholipase D catalytic domain (PLDC\_SF domain) (30).

**SNARE-associated proteins with phosphatase activity harbor a new motif.** In 19 out of 21 positive clones, we identified at least one gene encoding a protein domain



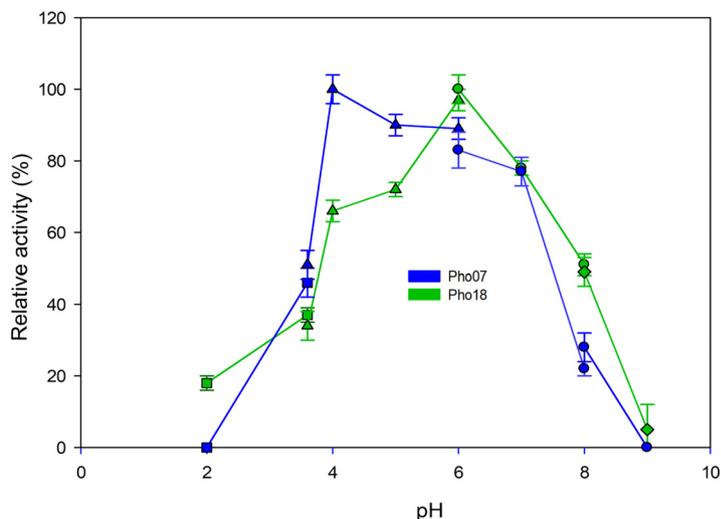
**FIG 2** Multiple sequence alignments of conserved regions of phosphatases belonging to the HP, PTP, and PAP2 superfamily. (A) Blue line, typical conserved HP phytase motif (RHGXRXP) in AEI69378 (phytase from *Yersinia mollaretii*) and AHA61669 (histidine acid phosphatase phytase from *Thermothelomyces thermophila*). Black line, the variations of the motif found in this study. (B) Typical PTP motif (HCX5R) in Pho14A, Pho16B, AAQ13669 (myo-inositol hexaphosphate phosphohydrolase from *Selenomonas ruminantium*), CAE79111 (protein tyrosine phosphatase 2 from *Bdellovibrio bacteriovorus* HD100), and CDI64125 (protein tyrosine phosphatase from *Xylophilus ampelinus*). (C) Catalytic sites of the PAP2 superfamily (cl00474), in Pho18, Pho24, CAI37740 (putative phosphatase from *Corynebacterium jeikeium*), NP\_639570 (phosphatase from *Xanthomonas campestris*), and BAC52270 (phosphatase from *Bradyrhizobium diazoefficiens*).

associated with catalytic activity of phosphatases. In contrast, the phosphatase-related genes of plasmids pLP04 and pLP15 did not encode known catalytic domains or signatures directly or indirectly associated with phosphatases. Clones carrying these plasmids showed significant phosphatase activity, and the products Pho04 and Pho15 showed sequence similarity to other previously reported proteins carrying the SNARE domain. However, both proteins shared overall sequence identity to previously reported phosphatases (Table 2). After individual heterologous expression of *pho04* and *pho15*, phosphatase activity was confirmed for both gene products. Pho04 and Pho15 hold the SNARE-associated domain DedA. SNARE-associated proteins are classified as structural proteins that function as a protein-protein interaction module (31). To our knowledge, no proteins with SNARE domains have been previously discovered to possess phosphatase activity.

We performed an alignment based on the *pho04* and *pho15* gene products, which revealed a shared conserved region (Fig. 3). Next, we analyzed all 56,539 sequences associated with the SNARE-associated Golgi proteins InterPro entry (IPR032816) with respect to motifs that were similar to those found in Pho04 and Pho15. A total of 905 sequences showed the conserved sequence pattern or a similar form. The sequence analysis revealed that Pho04 and Pho15 and the other 905 SNARE-associated (IPR032816) sequences share the particular amino acid arrangement ESSF(F/L/I/V)P. Notably, with respect to all analyzed proteins the identified motif was mostly from bacteria and detected outside the SNARE domain (cl00429) (examples are depicted in Fig. 3). Pho04 harbors the SNARE domain but shows 48% sequence identity to a putative membrane-associated alkaline phosphatase from *Acetobacter tropicalis*, while the closest phosphatase-related hit for Pho15 was an alkaline phosphatase from an *Acidobacteria* representative (43% identity) (Table 2).

**ALP-like superfamily and non-plant-derived PAP representatives showing phytase activity.** We selected the gene products of *pho07* and *pho18* for comprehensive

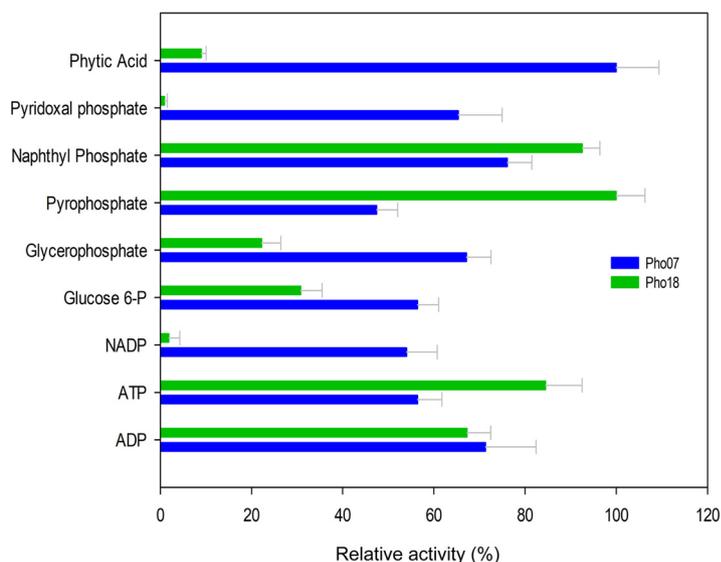




**FIG 5** Effect of pH on the relative activity of Pho07 and Pho18. The measurements were performed with different buffer systems according to the phytase standard assay at the optimal temperature of each protein. The average from triplicate experiments is presented. Glycine-HCl buffer, squares; sodium acetate buffer, triangles; Tris-maleate buffer, circles; glycine-NaOH buffer, diamonds. 100% relative phytase activity represented 4.84 and 1.39 U/mg for Pho07 and Pho18, respectively.

approximately 80% activity after incubation for 6 h at 40°C but lost more than 50% of its activity at temperatures  $\geq 50^\circ\text{C}$  (Fig. S4).

We evaluated the optimal pH range using different buffer systems at 30°C for Pho07 and at 50°C for Pho18. Pho07 exhibited the highest activity at pH 4.0 (Fig. 5) and retained more than 80% of its activity between pH 5.0 and 7.0. Low or no enzymatic activity was detected at pH values lower than 2.0 and higher than 8.0. Pho18 showed the highest activity at pH 6.0 and retained more than 70% of its activity at pH 5.0 and 7.0 (Fig. 5). To determine the substrate specificity of Pho07 and Pho18, we tested several phosphorylated compounds as the substrates (Fig. 6). Pho07 released phosphate from all tested compounds with the highest activity toward phytate and lowest activity toward pyrophosphate. Pho18 showed the highest relative activity with pyro-



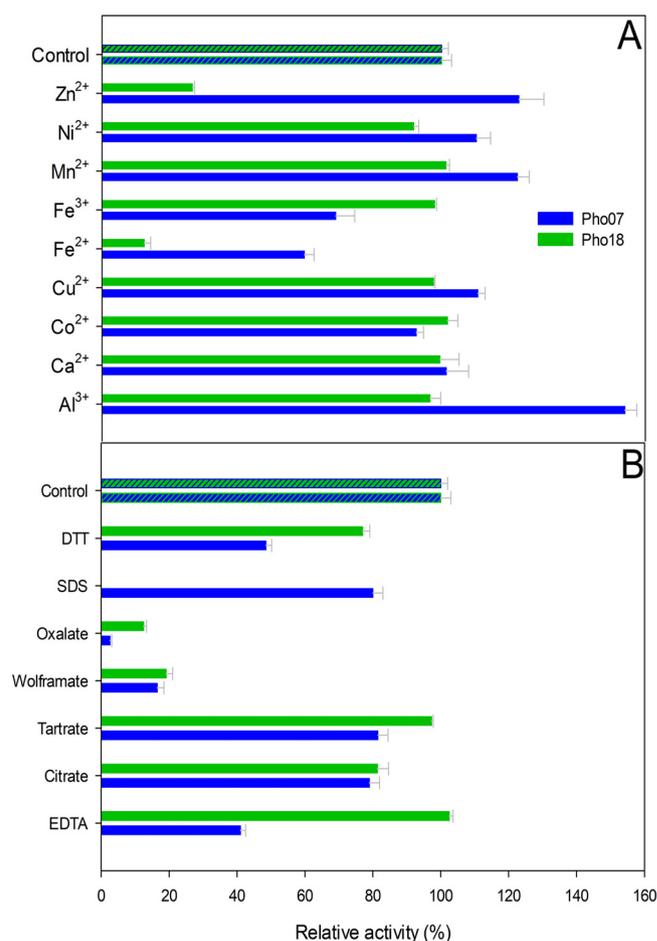
**FIG 6** Substrate specificity of Pho07 and Pho18. Specific activities corresponding to 100% relative phytase and pyrophosphatase activities of Pho07 and Pho18 were 2.98 and 13.3 U/mg, respectively. All measurements were performed in triplicate and under optimal pH and temperature conditions for each enzyme.

**TABLE 3** Kinetic values of Pho07 and Pho18 under optimal pH and temperature conditions

Enzyme	Mean (3 expts) ± SD							
	$K_m$ (mM)		$V_{max}$ ( $\mu\text{mol min}^{-1} \text{mg}^{-1}$ )		$k_{cat}$ ( $\text{min}^{-1}$ )		$k_{cat}/K_m$ ( $\text{min}^{-1} \text{M}^{-1}$ )	
	Sodium phytate	Pyrophosphate	Sodium phytate	Pyrophosphate	Sodium phytate	Pyrophosphate	Sodium phytate	Pyrophosphate
Pho07	0.49 ± 0.18	1.09 ± 0.03	6.50E-03 ± 1.01E-06	1.30E-04 ± 8.05E-06	694 ± 12.43	516 ± 22.98	3,410 ± 122	4,991 ± 155
Pho18	0.96 ± 0.09	0.22 ± 0.04	2.82E-03 ± 2.01E-04	4.03E-04 ± 4.42E-07	152 ± 9.83	1,088 ± 34.09	1,550 ± 18	49,200 ± 274

phosphate as the substrate and no significant activity with pyridoxal phosphate and NADP. As Pho07 and Pho18 exhibited the highest activity with phytate and pyrophosphate, respectively, we used these substrates for calculation of kinetic constants (Table 3).

Finally, we measured the effect of various metal ions and potential enzyme inhibitors on the activity of Pho07 and Pho18 with phytate as the substrate (Fig. 7). The metal ions showed different effects on the activity of the analyzed proteins.  $\text{Al}^{3+}$ ,  $\text{Mn}^{2+}$ , and  $\text{Zn}^{2+}$  increased the activity of Pho07, while the activity of Pho18 decreased in the presence of  $\text{Zn}^{2+}$ .  $\text{Fe}^{2+}$  had a strong inhibitory effect on the activity of both enzymes. With respect to potential inhibitors, the strongest inhibitory effects were observed at concentrations of 1 mM. Pho07 and Pho18 activities were reduced by most of the



**FIG 7** Effect of (A) metal ions and (B) potential inhibitors at 1 mM on the relative activity of Pho07 and Pho18. Specific activity values expressed as percentages of the control reactions are 3.8 and 1.3 U/mg for Pho07 and Pho18, respectively (A), and 3.5 and 1.22 U/mg for Pho07 and Pho18 (B), respectively.

tested inhibitors. Oxalate was the strongest inhibitor for Pho07, while the activity of Pho18 was completely depleted in the presence of SDS (Fig. 7).

## DISCUSSION

Apart from the relationship with pathogenesis and the economic relevance, phosphatases/phytases are also important for reduction of phosphorus pollution and its impact on diverse environments (8, 11, 13). However, only a few phosphatases, most of them from cultivable organisms, have been comprehensively analyzed. The discovery of new phosphatases from environmental samples as well as engineering of available representatives of this enzyme group is considered a major research challenge (33). So far, few studies have attempted to discover phosphatases/phytases encoded by metagenomes using a function-based approach. Within these studies, only three genes and one of the corresponding proteins which exhibited phytase activity were recovered and described (34–36). We found 31 candidate genes, and 24 of them encoded phosphatase activity after individual heterologous expression (Table 2). For the remaining seven genes, activity was not detected at individual gene level. The corresponding gene products might be part of larger phosphatase units or require other components encoded by the insert to show phosphatase activity.

Approximately 55% of the gene products described in this study showed low protein sequence identity to known phosphatases (50% or less) (Table 2), which demonstrates the capacity of our screening method to identify novel enzymes with phosphatase activity from environmental samples. It has been previously discovered that the absence of free phosphate and the addition of phytate to medium induce the expression of phytases (37). Therefore, it is indicated that many of the detected genes encode new enzymes with phytase activity as observed for Pho07 and Pho18.

ALP phosphomonoesterases widely occur in nature. They preferably hydrolyze phosphate esters at pH levels higher than 7.0 (38). The ALP-like superfamily (cl23718) was the most abundant domain we detected in the recovered hits derived from our soil metagenomic libraries. The pH of the soil samples used ranged from 3.1 to 4.5 (39). Nevertheless, acid phosphatase genes are considered to be more abundant than alkaline phosphatase genes in low-pH soils. This might be due to the fact that most studies on the prevalence of alkaline and acid phosphatase genes are based on PCR-based gene amplification using specific known genes from cultured individual species as starting point for primer design (40). This approach covers only a small fraction of the existent functional phosphatase genes. Here, we revealed the existence of so-far-unknown functional ALPs with low identity toward known phosphatases, evidencing the potential of our functional metagenomic approach for the discovery of new ALP-phosphatases from environmental samples.

To our knowledge, enzymes from the ALP-like superfamily entry (cl23718) exhibiting phytase activity have not been described or comprehensively characterized yet. Nevertheless, numerous proteins are mentioned in literature or annotated in databases as alkaline phosphatases with phytase activity, but their molecular signatures and domains are associated mostly with the classic phytases (14). The analysis by Lim et al. (41) focusing on the distribution and diversity of phytate-mineralizing bacteria considers alkaline phosphatases to be ubiquitous in living organisms and shows that they dephosphorylate a wide range of P compounds, but not phytate. Thus, the functional proteins carrying the ALP-like superfamily domain reported in this study (7) represent a new group of phytase enzymes. The phylogenetic analysis of the ALP-like members revealed that most of our metagenome-derived enzymes cluster separately from previously reported alkaline phosphatases/phytases (see Fig. S2 in the supplemental material).

The biochemical analysis of a selected ALP-like member, Pho07, showed that its temperature optimum is similar to the metagenome-derived alkaline phosphatase (mAP). This enzyme is one of the few reported phosphatases derived from environmental samples and not associated with cultures (42). Furthermore, the optimal pH range of Pho07 (4.0 to 5.0) is similar to that of other soil bacterial phytases (43). Among

the tested substrates, Pho07 showed the highest activity toward phytate, indicating that its primary activity is related to the degradation of this compound. Several studies report an enhancing effect of  $\text{Ca}^{2+}$  and  $\text{Mn}^{2+}$  on phytase activity (43). Nevertheless, the activity of Pho07 increased in the presence of  $\text{Mn}^{2+}$ , but it was not affected by  $\text{Ca}^{2+}$ . Among the potential inhibitors, wolframate and oxalate did not show significant effects on the activity of a phytate-degrading enzyme from *Pantoea agglomerans* (44) but reduced the relative activity of Pho07 to values lower than 20%. Since Pho07 is the first reported phytase carrying an ALP-like domain, it is not possible to compare its kinetic parameters (Table 3) with those from phytases of the same type.

The enzyme Pho18 belongs to the known PAPphy group of phytases. Only a few examples of characterized PAP proteins with phytase activity have been previously reported, and all of them were derived from plants (25). However, the presence of PAP-related genes in mammals, fungi, and bacteria has been indicated based on annotated genome sequences. The taxonomic analysis of *pho18* and the complete insert harboring it revealed a bacterial origin and a phylogenetic association with the genus *Terrimicrobium* of the *Verrucomicrobia* phylum (Table S1). In addition, biochemical analysis confirmed phytase activity of Pho18. Therefore, we report here for the first time a PAP2 phosphatase with phytase activity, which is of nonplant origin and metagenome derived. Moreover, the phylogenetic analysis showed that Pho18 clusters separately from other previously reported PAPs with phytase activity. The reason for this is most likely the vegetal origin of the previously reported PAP phytases (Fig. S2). To our knowledge, the study of Ghorbani Nasrabadi et al. (45) is the only attempt to identify PAP phytases derived from bacteria. In their study, an indirect association between phytase activity and the amplification of a putative PAP gene in the bacterial host was established (45).

The optimal temperature of Pho18 (50°C) is similar to optimal temperatures of other PAPs derived from wheat (45°C) and soybeans (58°C) (14). Furthermore, the behavior of Pho18 at temperatures higher than 55°C (Fig. 4) is similar to that reported for soybean phytases (46). An increase of phytase activity mediated by the addition of  $\text{Mn}^{2+}$  was reported for PAP phytases (32, 43). We did not register significant increases in the activity of Pho18 in the presence of any cation. However, the enzyme was strongly inhibited by  $\text{Zn}^{2+}$ , which is in contrast to other PAP phytases showing higher activity in the presence of this ion. Although Pho18 exhibits higher affinity to pyrophosphate, the kinetic parameters using phytate as the substrate are similar to PAP phytases from *Arabidopsis* (Table 3) (47).

We found the HAD (cl21460) domain as the second most abundant domain in our survey. The HAD domain is present in proteins of diverse organisms, including bacteria, archaea, and eukaryotes (48). This domain is carried by proteins able to catalyze a variety of biological functions and act on a wide range of substrates (19). Numerous members of the HAD superfamily can transfer phosphoryl groups or act as phosphoanhydride hydrolase P-type ATPases (49). Since proteins harboring this domain are involved in a variety of cellular processes, it is not surprising that they can be isolated through functional metagenomic screening for phosphatases.

One of the most remarkable findings in this study was the detection of the SNARE-associated domain (DedA, InterPro entry IPR032816) of Pho04 and Pho15. So far, the role of the SNARE-associated domain (DedA) has not been deeply studied. Bacterial DedA family mutants display phenotypes evidencing cell division defects, temperature sensitivity, and altered membrane phospholipid composition among others (50). DedA-SNAREs have been reported to promote or block membrane fusion, particularly during bacterial pathogenic processes (51). To our knowledge no phosphatase activity has been reported for proteins harboring SNARE-associated domains. Moreover, the particular signature ESSF(F/L/I/V)P has been overlooked until now.

In conclusion, we demonstrate here for the first time the potential of functional metagenomics to exploit the phosphatase pools hidden in environmental samples. Our study revealed new phosphatases/phytases with diverse and, so far, largely unknown characteristics. Furthermore, we discovered the existence of a new type of phytases

(ALP-like-phy) and found that the classical PAPphy are also functional in microorganisms and not only in plants.

## MATERIALS AND METHODS

**Soil sampling, DNA extraction, and construction of metagenomic libraries.** Genes encoding phosphatases were recovered from metagenomic libraries derived from A horizons of soil samples, which had been taken from forest sites of the German Biodiversity Exploratories Schwäbische Alb (samples AEW1 and AEW5), Hainich-Dün (sample HEW30), and Schorfheide-Chorin (samples SEW2, SEW5, and SEW46). Collection of samples was performed previously as described by Kaiser et al. (52) and Nacke et al. (39), respectively. Soil characteristics are available in Nacke et al. (39). Names of constructed metagenomic libraries refer to the designation of the samples from which the libraries were derived. Metagenomic libraries were generated using the method described by Nacke et al. (39). The plasmid libraries AEW1, AEW5, SEW2, and SEW5 have been previously generated by employing the same approach (39).

**Function-based screening and identification of ORFs encoding phosphatase activity.** For function-based screening of metagenomic libraries, we used our recently described method (18). Small-insert libraries were constructed using the plasmid pCR-XL-TOPO as vector (Invitrogen GmbH, Karlsruhe, Germany) and *Escherichia coli* DH5 $\alpha$  [F<sup>-</sup>  $\phi$ 80lacZ $\Delta$ M15  $\Delta$ (lacZYA-argF)U169 recA1 endA1 hsdR17(r<sub>K</sub><sup>-</sup> m<sub>K</sub><sup>+</sup>) phoA supE44  $\lambda$ -thi-1 gyrA96 relA1] as screening host. Modified Sperber medium (16 g/liter agar, 10 g/liter glucose, 500 mg/liter yeast extract, 100 mg/liter CaCl<sub>2</sub>, and 250 mg/liter MgSO<sub>4</sub>) was used as a screening medium supplemented with 2.5 g/liter phytic acid as sole P source and 25  $\mu$ g/ml of 5-bromo-4-chloro-3-indolyl phosphate (BCIP) (53). The modified Sperber minimal medium used in this study was used for detection of phosphatase/phytase activity of the library-bearing *E. coli* clones due to the presence of phytate and the absence of other inorganic P sources (37). The slight background activity observed after more than 48 h of incubation of the host strain is probably caused by the alkaline phosphatase-encoding gene (*phoA*) of the host. Positive clones show an intense dark blue colony color, whereas negative colonies exhibit first a white and subsequently a light blue or green color after prolonged incubation.

The plasmids derived from positive clones were sequenced by the Göttingen Genomics Laboratory (Göttingen, Germany), and ORF prediction was performed as described by Nacke et al. (39). Next, the obtained sequences were analyzed by using the Basic Local Alignment Search Tool (BLAST) (54). Only plasmids harboring at least one ORF potentially associated with phosphatase activity were considered candidates for further analysis and designated pLP01 to pLP04, pLP07 to pLP10, pLP13 to pLP20, and pLP24 to pLP28. Full-length sequence alignment was performed between the candidates and their closest related sequence by using Clustal Omega (55). All coding sequences were examined for similarities to known protein families and domains by performing searches against the InterPro collection of protein signature databases and conserved domain databases (CDD) (56, 57). The prediction of signal peptides of the proteins was performed by using SignalP 4.0 (58). Additionally, all inserts were taxonomically classified by using the software Kaiju 1.5.0 (59). Alignments of the deduced protein sequences and phylogenetic trees of the proteins were performed by using MEGA 7 (60). The maximum likelihood method based on the equal input model was applied. The bootstrap values were calculated from 500 replicates, and branches corresponding to partitions reproduced in fewer than 50% of bootstrap replicates were collapsed. Alignments were visualized by using Jalview version 2 (61).

Candidate genes encoding domains that have not been previously associated with phosphatase activity (pLP04 and pLP15) and inserts comprising more than one potential phosphatase-encoding gene were amplified and subsequently cloned. Specific primers for each target gene were designed, and the pET101/D directional TOPO cloning kit was used for cloning as recommended by the manufacturer (Thermo Fisher Scientific GmbH, Schwerte, Germany). PCR was carried out in a 50- $\mu$ l volume containing 10  $\mu$ l of 5-fold Phusion GC buffer, 200  $\mu$ M (each) dNTP, 1.5 mM MgCl<sub>2</sub>, 2  $\mu$ M (each) primers, 2.5% DMSO, 0.5 U Phusion High Fidelity Hot Start DNA polymerase (Thermo Fisher Scientific GmbH, Schwerte, Germany), and 25 ng recombinant plasmid. PCR conditions were as follows: initial denaturation at 98°C for 2 min followed by 30 cycles of denaturation at 98°C for 1.5 min, annealing at 58°C for 1 min, and extension at 72°C for 1 min, followed by a final extension at 72°C for 5 min. Subsequently, the amplified genes were individually cloned into the expression vector pET 101/D and transformed into *E. coli* BL21 [F<sup>-</sup> *ompT hsdS<sub>B</sub>*(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) *gal dcm* (DE3)] as recommended by the manufacturer (Thermo Fisher Scientific GmbH). The resulting recombinant plasmid-bearing *E. coli* BL21 strains were subsequently plated on Sperber minimal medium agar supplemented with phytic acid (2.5 g/liter), BCIP (25 mg/ml), and IPTG (0.25 mM) for phosphatase activity detection.

**Heterologous expression of *pho07* and *pho18* and purification of gene products.** The genes *pho07* and *pho18* carried by plasmids pLP07 and pLP18, respectively, were selected for heterologous expression with the pET-20b (+) (V5-epitope/His tag) vector (Merck KGaA, Darmstadt, Germany) as recommended by Villamizar et al. for metagenome-derived phosphatases (18). Crude extracts containing the target proteins were derived from the expression strain *E. coli* BL21 and filtered as described by Villamizar et al. (18). For purification of the proteins, the filtered crude extracts were then transferred to nickel columns (Protino2000 Ni-Ted columns; Macherey and Nagel, Düren, Germany). The equilibration of the columns and the washing steps were performed with 50 mM HEPES buffer (pH 8.0) containing 200 mM NaCl, followed by three elution steps with 50 mM HEPES, 200 mM NaCl, and 250 mM imidazole. Pho07 was further purified by using the Äkta FPLC system (GE Healthcare, Little Chalfont, United Kingdom) via hydrophobic interaction chromatography. A 15PHE 4.6/100PE Tricorn high-performance column (GE Healthcare, Little Chalfont, United Kingdom) in a total bed volume of 1.7 ml with a 2-ml/min

flow rate at room temperature was utilized. Pho18 was purified through ion-exchange chromatography, by using a cation exchanger (SOURCE15S) in a prepacked Tricorn column (4.6/100 PE) (GE Healthcare, Little Chalfont, United Kingdom) with a gel bed volume of 1.7 ml at a 1-ml/min flow rate and room temperature. The purity of the resulting protein preparations was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and the detection of V5 epitope-carrying proteins was achieved by Western blot hybridization, as described by Waschkowitz et al. (62).

**Enzyme assays.** Phosphatase activity was determined at 355 nm by detecting the release of inorganic phosphorus according to the ammonium molybdate method developed by Heinonen and Lahti with modifications (44, 63) as follows: the enzyme solutions (10  $\mu$ l) were preincubated for 3 min at 40°C in 380  $\mu$ l of 50 mM sodium acetate buffer (pH 5). Subsequently, 10  $\mu$ l of 100 mM phytic acid dipotassium salt (Sigma-Aldrich, Munich, Germany) was added, and the mixture was incubated for 30 min at 40°C. To stop the reaction, 1.5 ml of freshly prepared AAM solution (acetone–5 N H<sub>2</sub>SO<sub>4</sub>–10 mM ammonium molybdate) and 100  $\mu$ l of 1 M citric acid were added. Blanks were prepared by adding AAM solution prior to the addition of enzyme. The absorbance (355 nm) was measured using the Ultrospec 3300 Pro (Amersham plc, Little Chalfont, United Kingdom).

To assess the influence of pH on purified enzymes, the activity was measured at 40°C in a pH range from 1 to 9. The following overlapping buffer systems were used: 50 mM glycine-HCl (pH 1.0 to 3.5), 50 mM sodium acetate (pH 3.5 to 6.0), 50 mM Tris-maleate acid (pH 6.0 to 8.0), and 50 mM glycine-NaOH (pH 7.0 to 9.0). After the optimal pH was determined for Pho07 and Pho18, the influence of temperature on enzymatic activity was analyzed. The thermal stability was checked after incubation of the purified enzymes at different temperatures.

The substrate specificity of the phosphatases was determined using the standard assay described above under the optimal temperature and pH for each enzyme (substrate concentration, 10 mM). Furthermore, the effects of cations (Al<sup>3+</sup>, Ca<sup>2+</sup>, Co<sup>2+</sup>, Fe<sup>2+</sup>, Fe<sup>3+</sup>, Mn<sup>2+</sup>, Ni<sup>2+</sup>, and Zn<sup>2+</sup>) and the potential inhibitors (EDTA, citrate, tartrate, wolframate, oxalate, sodium dodecyl sulfate (SDS), and dithiothreitol (DTT) at concentrations of 0.1 and 1 mM were analyzed.

For the kinetic constants, all measurements were performed in triplicate under optimal pH and temperature conditions using phytic acid and pyrophosphate as the substrates. The data were analyzed by the Sigma Plot Enzyme Kinetic Module version SigmaPlot 12.0 (Systat Software, Inc., San Jose, CA).

**Sequence accession numbers.** The nucleotide sequences of plasmids listed in Table 1 have been submitted to the National Center for Biotechnology Information (NCBI) GenBank database under the accession numbers indicated: pLP01 (Pho01), [KY931670](#); pLP02 (Pho02), [KY931671](#); pLP03 (Pho03A and -B), [KY931672](#); pLP04 (Pho04), [KY931673](#); pLP07 (Pho07), [KY931674](#); pLP08 (Pho08A to -C), [KY931675](#); pLP09 (Pho09C), [KY931676](#); pLP10 (Pho10), [KY931677](#); pLP13 (Pho13), [KY931678](#); pLP14 (Pho14A to -D), [KY931679](#); pLP15 (Pho15), [KY931680](#); pLP16 (Pho16A and -B), [KY931681](#); pLP17 (Pho17A), [KY931682](#); pLP18 (Pho18), [KY931683](#); pLP19 (Pho19A), [KY931684](#); pLP20 (Pho20B), [KY931685](#); pLP24 (Pho24), [KY931686](#); pLP25 (Pho25B and -C), [KY931687](#); pLP26 (Pho26), [KY931688](#); pLP27 (Pho27A and -B), [KY931689](#); and pLP28 (Pho28A and Pho28C), [KY931690](#).

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mBio.01966-18>.

**FIG S1**, PDF file, 0.5 MB.

**FIG S2**, PDF file, 1.1 MB.

**FIG S3**, PDF file, 0.1 MB.

**FIG S4**, PDF file, 0.2 MB.

**TABLE S1**, PDF file, 0.04 MB.

## ACKNOWLEDGMENTS

We thank the Deutscher Akademischer Austauschdienst (DAAD) and Colciencias Colombia for the financial support to Genis Andrés Castillo in the frame of the fellowship program ALECOL. Additionally, we acknowledge support by the DFG and the Open Access Publication Funds of the Göttingen University.

We thank Silja Brady and Mechthild Bömeke for assistance with respect to expression and purification of the proteins, as well as Dominik Schneider for supporting the taxonomic classification of the inserts. We thank Elisabeth Gullans for her contribution to metagenomic library screening.

## REFERENCES

- Daniel R. 2005. The metagenomics of soil. *Nat Rev Microbiol* 3:470–478. <https://doi.org/10.1038/nrmicro1160>.
- Nacke H, Daniel R. 2014. Approaches in metagenome research: progress and challenges, p 38–43. In Nelson KE (ed), *Encyclopedia of metagenomics*. Springer, New York, NY.
- Adam N, Perner M. 2018. Novel hydrogenases from deep-sea hydrothermal vent metagenomes identified by a recently developed activity-based screen. *ISME J* 12:1225–1236. <https://doi.org/10.1038/s41396-017-0040-6>.
- Berini F, Casciello C, Marcone GL, Marinelli F. 2017. Metagenomics: novel

- enzymes from non-culturable microbes. *FEMS Microbiol Lett* 364:fnx211. <https://doi.org/10.1093/femsle/fnx211>.
5. Blüher D, Laha D, Thieme S, Hofer A, Eschen-Lippold L, Masch A, Balcke G, Pavlovic I, Nagel O, Schonsky A, Hinkelmann R, Worner J, Parvin N, Greiner R, Weber S, Tissier A, Schutkowski M, Lee J, Jessen H, Schaaf G, Bonas U. 2017. A 1-phytase type III effector interferes with plant hormone signaling. *Nat Commun* 8:2159. <https://doi.org/10.1038/s41467-017-02195-8>.
  6. Chen MJ, Dixon JE, Manning G. 2017. Genomics and evolution of protein phosphatases. *Sci Signal* 10:eaag1796. <https://doi.org/10.1126/scisignal.aag1796>.
  7. Golovan SP, Meidinger RG, Ajakaiye A, Cottrill M, Wiederkehr MZ, Barney DJ, Plante C, Pollard JW, Fan MZ, Hayes MA, Laursen J, Hjorth JP, Hacker RR, Phillips JP, Forsberg CW. 2001. Pigs expressing salivary phytase produce low-phosphorus manure. *Nat Biotechnol* 19:741–745. <https://doi.org/10.1038/90788>.
  8. Elser J, Bennett E. 2011. Phosphorus cycle: a broken biogeochemical cycle. *Nature* 478:29–31. <https://doi.org/10.1038/478029a>.
  9. Naylor RL, Hardy RW, Bureau DP, Chiu A, Elliott M, Farrell AP, Forster I, Gatlin DM, Goldberg RJ, Hua K, Nichols PD. 2009. Feeding aquaculture in an era of finite resources. *Proc Natl Acad Sci U S A* 106:15103–15110. <https://doi.org/10.1073/pnas.0905235106>.
  10. Scholz RW, Ulrich AE, Eilittä M, Roy A. 2013. Sustainable use of phosphorus: a finite resource. *Sci Total Environ* 461-462:799–803. <https://doi.org/10.1016/j.scitotenv.2013.05.043>.
  11. Zhang X, Li Z, Yang H, Liu D, Cai G, Li G, Mo J, Wang D, Zhong C, Wang H, Sun Y, Shi J, Zheng E, Meng F, Zhang M, He X, Zhou R, Zhang J, Huang M, Zhang R, Li N, Fan M, Yang J, Wu Z. 2018. Novel transgenic pigs with enhanced growth and reduced environmental impact. *Elife* 7:e34286. <https://doi.org/10.7554/eLife.34286>.
  12. Schröder JJ, Smit AL, Cordell D, Rosemarin A. 2011. Improved phosphorus use efficiency in agriculture: a key requirement for its sustainable use. *Chemosphere* 84:822–831. <https://doi.org/10.1016/j.chemosphere.2011.01.065>.
  13. Ushasree MV, Vidya J, Pandey A. 2017. Other enzymes: phytases, p 309–333. In Pandey A, Negi S, Soccol CR (ed), *Current developments in biotechnology and bioengineering*. Elsevier, Amsterdam, Netherlands.
  14. Lei XG, Weaver JD, Mullaney E, Ullah AH, Azain MJ. 2013. Phytase, a new life for an “old” enzyme. *Annu Rev Anim Biosci* 1:283–309. <https://doi.org/10.1146/annurev-animal-031412-103717>.
  15. Jiang X, Sobetzko P, Nasser W, Reverchon S, Muskhelishvili G. 2015. Chromosomal “stress-response” domains govern the spatiotemporal expression of the bacterial virulence program. *mBio* 6:e00353-15. <https://doi.org/10.1128/mBio.00353-15>.
  16. Tsang PW, Fong WP, Samaranyake LP. 2017. *Candida albicans* orf19.3727 encodes phytase activity and is essential for human tissue damage. *PLoS One* 12:e0189219. <https://doi.org/10.1371/journal.pone.0189219>.
  17. Mullaney EJ, Ullah AHJ. 2003. The term phytase comprises several different classes of enzymes. *Biochem Biophys Res Commun* 312: 179–184. <https://doi.org/10.1016/j.bbrc.2003.09.176>.
  18. Villamizar GAC, Nacke H, Daniel R. 2017. Function-based metagenomic library screening and heterologous expression strategy for genes encoding phosphatase activity. *Methods Mol Biol* 1539:249–260. [https://doi.org/10.1007/978-1-4939-6691-2\\_16](https://doi.org/10.1007/978-1-4939-6691-2_16).
  19. Huang H, Pandya C, Liu C, Al-Obaidi NF, Wang M, Zheng L, Toews Keating S, Aono M, Love JD, Evans B, Seidel RD, Hillerich BS, Garforth SJ, Almo SC, Mariano PS, Dunaway-Mariano D, Allen KN, Farelli JD. 2015. Panoramic view of a superfamily of phosphatases through substrate profiling. *Proc Natl Acad Sci U S A* 112:E1974–E1983. <https://doi.org/10.1073/pnas.1423570112>.
  20. Mustelin T. 2007. A brief introduction to the protein phosphatase families, p 9–22. In Moorhead G (ed), *Protein phosphatase protocols*. Springer, Totowa, NJ.
  21. Rigden DJ. 2008. The histidine phosphatase superfamily: structure and function. *Biochem J* 409:333–348. <https://doi.org/10.1042/BJ20071097>.
  22. Oh BC, Choi WC, Park S, Kim YO, Oh TK. 2004. Biochemical properties and substrate specificities of alkaline and histidine acid phytases. *Appl Microbiol Biotechnol* 63:362–372. <https://doi.org/10.1007/s00253-003-1345-0>.
  23. Puhl AA, Gruninger RJ, Greiner R, Janzen TW, Mosimann SC, Selinger LB. 2007. Kinetic and structural analysis of a bacterial protein tyrosine phosphatase-like myo-inositol polyphosphatase. *Protein Sci* 16:1368–1378. <https://doi.org/10.1110/ps.062738307>.
  24. Gruninger RJ, Dobing S, Smith AD, Bruder LM, Selinger LB, Wieden HJ, Mosimann SC. 2012. Substrate binding in protein-tyrosine phosphatase-like inositol polyphosphatases. *J Biol Chem* 287:9722–9730. <https://doi.org/10.1074/jbc.M111.309872>.
  25. Schenk G, Mitić N, Hanson GR, Comba P. 2013. Purple acid phosphatase: a journey into the function and mechanism of a colorful enzyme. *Coord Chem Rev* 257:473–482. <https://doi.org/10.1016/j.ccr.2012.03.020>.
  26. Nardini M, Dijkstra BW. 1999.  $\alpha/\beta$  hydrolase fold enzymes: the family keeps growing. *Curr Opin Struct Biol* 9:732–737. [https://doi.org/10.1016/S0959-440X\(99\)00037-8](https://doi.org/10.1016/S0959-440X(99)00037-8).
  27. Holmquist M. 2000. Alpha beta-hydrolase fold enzymes structures, functions and mechanisms. *Curr Protein Pept Sci* 1:209–235. <https://doi.org/10.2174/1389203003381405>.
  28. Simm R, Morr M, Kader A, Nimitz M, Römmling U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* 53:1123–1134. <https://doi.org/10.1111/j.1365-2958.2004.04206.x>.
  29. Leipe DD, Koonin EV, Aravind L. 2004. STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. *J Mol Biol* 343:1–28. <https://doi.org/10.1016/j.jmb.2004.08.023>.
  30. Waite M. 1999. The PLD superfamily: insights into catalysis. *Biochim Biophys Acta* 1439:187–197. [https://doi.org/10.1016/S1388-1981\(99\)00094-3](https://doi.org/10.1016/S1388-1981(99)00094-3).
  31. Fratti RA, Wickner W. 2007. Distinct targeting and fusion functions of the PX and SNARE domains of yeast vacuolar Vam7p. *J Biol Chem* 282: 13133–13138. <https://doi.org/10.1074/jbc.M700584200>.
  32. Dionisio G, Madsen CK, Holm PB, Welinder KG, Jorgensen M, Stoger E, Arcalis E, Brinch-Pedersen H. 2011. Cloning and characterization of purple acid phosphatase phytases from wheat, barley, maize, and rice. *Plant Physiol* 156:1087–1100. <https://doi.org/10.1104/pp.110.164756>.
  33. Kumar A, Chanderman A, Makolomakwa M, Perumal K, Singh S. 2016. Microbial production of phytases for combating environmental phosphate pollution and other diverse applications. *Crit Rev Environ Sci Technol* 46:556–591. <https://doi.org/10.1080/10643389.2015.1131562>.
  34. Tan H, Mooij MJ, Barret M, Hegarty PM, Harington C, Dobson AD, O’Gara F. 2014. Identification of novel phytase genes from an agricultural soil-derived metagenome. *J Microbiol Biotechnol* 24:113–118. <https://doi.org/10.4014/jmb.1307.07007>.
  35. Lee MH, Oh KH, Kang CH, Kim JH, Oh TK, Ryu CM, Yoon JH. 2012. Novel metagenome-derived, cold-adapted alkaline phospholipase with superior lipase activity as an intermediate between phospholipase and lipase. *Appl Environ Microbiol* 78:4959–4966. <https://doi.org/10.1128/AEM.00260-12>.
  36. Farias N, Almeida I, Meneses C. 2018. New bacterial phytase through metagenomic prospecting. *Molecules* 23:448. <https://doi.org/10.3390/molecules23020448>.
  37. Kerovuoto J, Lauraeus M, Nurminen P, Kalkkinen N, Apajalahti J. 1998. Isolation, characterization, molecular gene cloning, and sequencing of a novel phytase from *Bacillus subtilis*. *Appl Environ Microbiol* 64:2079–2085.
  38. Sharma U, Pal D, Prasad R. 2014. Alkaline phosphatase: an overview. *Indian J Clin Biochem* 29:269–278. <https://doi.org/10.1007/s12291-013-0408-y>.
  39. Nacke H, Will C, Herzog S, Nowka B, Engelhaupt M, Daniel R. 2011. Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories. *FEMS Microbiol Ecol* 78:188–201. <https://doi.org/10.1111/j.1574-6941.2011.01088.x>.
  40. Fraser T, Lynch DH, Entz MH, Dunfield KE. 2015. Linking alkaline phosphatase activity with bacterial phoD gene abundance in soil from a long-term management trial. *Geoderma* 257:115–122. <https://doi.org/10.1016/j.geoderma.2014.10.016>.
  41. Lim BL, Yeung P, Cheng C, Hill JE. 2007. Distribution and diversity of phytate-mineralizing bacteria. *ISME J* 1:321–330. <https://doi.org/10.1038/ismej.2007.40>.
  42. Lee DH, Choi SL, Rha E, Kim SJ, Yeom SJ, Moon JH, Lee SG. 2015. A novel psychrophilic alkaline phosphatase from the metagenome of tidal flat sediments. *BMC Biotechnol* 15:1. <https://doi.org/10.1186/s12896-015-0115-2>.
  43. Suleimanova AD, Beinbauer A, Valeeva LR, Chastukhina IB, Balaban NP, Shakirov EV, Greiner R, Sharipova MR. 2015. Novel glucose-1-phosphatase with high phytase activity and unusual metal ion activation from soil

- bacterium *Pantoea* sp. strain 3.5.1. *Appl Environ Microbiol* 81: 6790–6799. <https://doi.org/10.1128/AEM.01384-15>.
44. Greiner R. 2004. Purification and properties of a phytate-degrading enzyme from *Pantoea agglomerans*. *Protein J* 23:567–576. <https://doi.org/10.1007/s10930-004-7883-1>.
  45. Ghorbani Nasrabadi R, Greiner R, Yamchi A, Nourzadeh Roshan E. 2018. A novel purple acid phytase from an earthworm cast bacterium. *J Sci Food Agric* 98:3667–3674. <https://doi.org/10.1002/jsfa.8845>.
  46. Hegeman CE, Grabau EA. 2001. A novel phytase with sequence similarity to purple acid phosphatases is expressed in cotyledons of germinating soybean seedlings. *Plant Physiol* 126:1598–1608. <https://doi.org/10.1104/pp.126.4.1598>.
  47. Kuang R, Chan KH, Yeung E, Lim BL. 2009. Molecular and biochemical characterization of AtPAP15, a purple acid phosphatase with phytase activity, in *Arabidopsis*. *Plant Physiol* 151:199–209. <https://doi.org/10.1104/pp.109.143180>.
  48. Burroughs AM, Allen KN, Dunaway-Mariano D, Aravind L. 2006. Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. *J Mol Biol* 361:1003–1034. <https://doi.org/10.1016/j.jmb.2006.06.049>.
  49. Koonin EV, Tatusov RL. 1994. Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. Application of an iterative approach to database search. *J Mol Biol* 244:125–132. <https://doi.org/10.1006/jmbi.1994.1711>.
  50. Doerrler WT, Sikdar R, Kumar S, Boughner LA. 2013. New functions for the ancient DedA membrane protein family. *J Bacteriol* 195:3–11. <https://doi.org/10.1128/JB.01006-12>.
  51. Wesolowski J, Paumet F. 2010. SNARE motif: a common motif used by pathogens to manipulate membrane fusion. *Virulence* 1:319–324. <https://doi.org/10.4161/viru.1.4.12195>.
  52. Kaiser K, Wemheuer B, Korolkow V, Wemheuer F, Nacke H, Schöning I, Schruppf M, Daniel R. 2016. Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. *Sci Rep* 6:33696. <https://doi.org/10.1038/srep33696>.
  53. Sperber JI. 1958. The incidence of apatite-solubilizing organisms in the rhizosphere and soil. *Aust J Agric Res* 9:778–781. <https://doi.org/10.1071/AR9580778>.
  54. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
  55. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539. <https://doi.org/10.1038/msb.2011.75>.
  56. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30: 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>.
  57. Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Zhang D, Bryant SH. 2013. CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res* 41:D348–D352. <https://doi.org/10.1093/nar/gks1243>.
  58. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785–786. <https://doi.org/10.1038/nmeth.1701>.
  59. Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7:11257. <https://doi.org/10.1038/ncomms11257>.
  60. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33: 1870–1874. <https://doi.org/10.1093/molbev/msw054>.
  61. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>.
  62. Waschkowitz T, Rockstroh S, Daniel R. 2009. Isolation and characterization of metalloproteases with a novel domain structure by construction and screening of metagenomic libraries. *Appl Environ Microbiol* 75: 2506–2516. <https://doi.org/10.1128/AEM.02136-08>.
  63. Heinonen JK, Lahti RJ. 1981. A new and convenient colorimetric determination of inorganic orthophosphate and its application to the assay of inorganic pyrophosphatase. *Anal Biochem* 113:313–317. [https://doi.org/10.1016/0003-2697\(81\)90082-8](https://doi.org/10.1016/0003-2697(81)90082-8).