J. Microbiol. Biotechnol. (2014), 24(1), 113–118
http://dx.doi.org/10.4014/jmb.1307.07007

jmb

# Identification of Novel Phytase Genes from an Agricultural Soil-Derived Metagenome

**Hao Tan[1], Marlies J. Mooij[1], Matthieu Barret[1], Pardraig M. Hegarty[1], Catriona Harrington[1], Alan D.W. Dobson[2], and Fergal O'Gara[1,2,3]***

[1]*BIOMERIT Research Centre, Microbiology Department, University College Cork, National University of Ireland, Cork, Ireland*
[2]*Environmental Research Institute, University College Cork, National University of Ireland, Cork, Ireland*
[3]*Curtin University, School of Biomedical Sciences, Perth, WA, Australia*

*Received: July 8, 2013*
*Revised: October 1, 2013*
*Accepted: October 10, 2013*

First published online
October 22, 2013

**Corresponding author*
Phone: +353-21-4902646;
Fax: +353-21-4903101;
E-mail: f.ogara@ucc.ie

Supplementary data for this paper are available on-line only at http://jmb.or.kr.

pISSN 1017-7825, eISSN 1738-8872



Environmental microorganisms are emerging as an important source of new enzymes for wide-scale industrial application. In this study, novel phytase genes were identified from a soil microbial community. For this, a function-based screening approach was utilized for the identification of phytase activity in a metagenomic library derived from an agricultural soil. Two novel phytases were identified. Interestingly, one of these phytases is an unusual histidine acid phosphatase family phytase, as the conserved motif of the active site of PhyX possesses an additional amino acid residue. The second phytase belongs to a new type, which is encoded by multiple open reading frames (ORFs) and is different to all phytases known to date, which are encoded by a single ORF.

**Keywords:** Phytase, phytate, functional metagenomics, HAP

Livestock feeding requires an effective removal of phytate from animal diets in order to enhance nutritional utilization [24, 28]. For this, the animal diets are usually supplemented with phytase [14, 18, 19, 25], a hydrolytic enzyme that can be produced by a number of microorganisms [15, 21]. Phytases can also be used in sustainable agriculture to convert soil phytate into phosphate, which can be taken up by crops as natural phosphorus (P) fertilizer [23]. Nowadays, identification and engineering of phytases are still in demand for various industrial applications and conditions [10, 13, 26, 31, 34]. To date, a majority of the already identified phytases were obtained based on sequence similarity [4], for example by using PCR amplification with degenerate primers [11, 12, 33]. However, such approaches are less competent in identifying novel phytases with no or low homology to known phytases. In this study, a function-based strategy was used to identify phytase encoded by novel genes from a soil environment.

For this, a metagenomic library was screened based on an ability to utilize phytate as sole P source. This library was derived from the soil of a farm cultivated with winter wheat (*Triticum aestivum*) [3]. This soil was subjected to no P fertilization application over a long time period. Environmental DNA fragments, with an average length of 35 kb, were directly extracted from the soil and then cloned into a fosmid vector, pCC1FOS (Epicentre Biotechnologies, USA). The library was transformed into *E. coli* EPI300 (Epicentre Biotechnologies) as a host.

To identify candidate clones with strong phytate utilization capacity, a selective medium containing phytate as the sole P source was used. The medium was prepared as previously described by Unno *et al*. [29]. A total of 14,400 clones in the metagenomic library were stamped onto the surface of the phytate medium using the Q-Pix II robot (Genetix, UK). *E. coli* EPI300 with pCC1FOS empty vector was used as a negative control clone. All plates were incubated for 5 days at 30°C. Clones with strong growth were selected as phytase-producing candidates. The candidate clones were re-streaked

on phytate medium to confirm their phytate utilization capacity. Two out of the 28 phytase-producing candidates obtained (WFS_1 and WFS_2) demonstrated significantly stronger growth compared with others and were selected for further characterization.

In order to identify the putative phytase gene(s) in these two clones, the fosmid was extracted from each clone and subsequently sent for sequence analysis (Eurofins MWG Operon, Germany). The fosmid insert sequences were annotated using RAST (Rapid Annotation using Subsystem Technology) [1]. Interestingly, no open read frame (ORF) was annotated as a phytase gene in either fosmid. This suggested that these two clones possess potential phytase genes with no obvious homology to any known phytase gene. In order to identify these novel phytase genes, motif search using WAPAM [7] was performed to identify conserved motifs belonging to all known phytase families, including the histidine acid phosphatase (HAP) family (RHGXRXP, X for a random amino acid) [30], the beta propeller (BPP) family (beta-sheet motif) [8, 27], the cysteine-phytase (CPhy) family (HCXXGXXRT/S) [2], and the purple acid phosphatase (PAP) family (DXG, GDXXY, GNHD/E, VXXH, or GHXH) [9, 17]. Interestingly, ORF_6 in clone WFS_1 (Table 1) was identified to possess a putative conserved motif belonging to the active site of the HAP phytase family. The amino acid sequence of the identified conserved motif was RHGLRYYP, which has an additional amino acid at the sixth position, in comparison with the typical conserved motif (RHGXRXP). In addition to this motif, HAP phytases contain another essential conserved motif, which consists of two amino acids "HD" near its C-terminal. The deduced amino acid sequence of ORF_6 also contains this "HD" motif, which together with the RHGLRYYP motif suggests that ORF_6 (further referred to as *phyX*) is likely to encode an unusual type of HAP phytase.

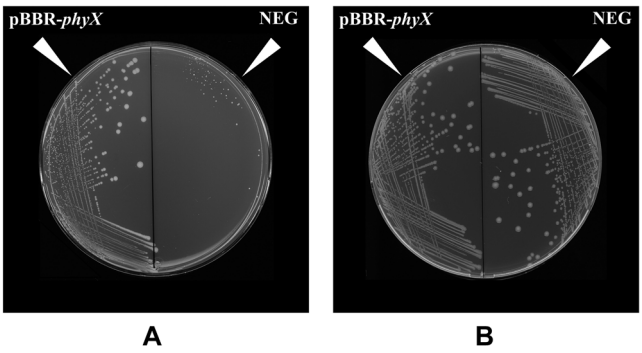


**Fig. 1.** Phytate utilization capacity of pBBR-*phyX*.
(**A**) Medium supplemented with 5 mg/ml phytate, and (**B**) medium supplemented with 10 mM inorganic phosphate as P source. The pBBR-*phyX* demonstrated a strong capacity to utilize phytate as a sole P source for growth, whereas the negative control clone (NEG, empty pBBR-MCS4 vector) had slight growth due to an inevitable background phytase activity contributed by the product of the *E. coli appA* gene. Replacement of phytate with inorganic phosphate showed no growth differences.

To confirm that the product of *phyX* is indeed a phytase, this ORF was subcloned into a pBBR1-MCS-4 vector [16] and expressed constitutively, as previously described by Patel *et al.* [23]. Primers *phyX*-F (5′-GAGCGAATTCACACG TAGAGGAGT-3′) and *phyX*-R (5′-ATGGTCTAGACTTCA AGCAGCCTTGA-3′) were used to amplify the *phyX* ORF

**Table 1.** An unusual HAP phytase (PhyX) encoded by ORF_6 of fosmid WFS_1, and its upstream and downstream ORFs. The annotated function of each ORF (using RAST) and amino acid (AA) sequence similarity to its closest hits (by BLASTP) are shown.

| ORF NO. in fosmid | Length of AA | Strand | Annotated function | Accession number of hit | Similarity of AA (%) | *E*-value | Taxonomic background of hit |
|---|---|---|---|---|---|---|---|
| 3 | 195 | + | Lysine decarboxylase family protein | YP_001021302 | 76 | 1E-47 | *Methylibium petroleiphilum* PM1 |
| 4 | 112 | - | Nitrogen regulatory protein P-II | YP_001791929 | 97 | 1E-49 | *Leptothrix cholodnii* SP-6 |
| 5 | 563 | - | NAD synthetase/ Glutamine amidotransferase chain of NAD synthetase | YP_001791995 | 80 | 0 | *Methylibium petroleiphilum* PM1 |
| 6 | 400 | + | hypothetical conserved protein | YP_001021305 | 72 | 3E-120 | *Methylibium petroleiphilum* PM1 |
| 7 | 902 | - | putative zinc protease | YP_001021310 | 72 | 0 | *Methylibium petroleiphilum* PM1 |
| 8 | 179 | - | Inorganic pyrophosphatase | YP_981936 | 85 | 4E-68 | *Polaromonas naphthalenivorans* CJ2 |
| 9 | 718 | + | Pyrophosphate-energized proton pump, V-Type h(+)-translocating pyrophosphatase | YP_981935 | 89 | 0 | *Geobacter uraniireducens* Rf4 |
| 10 | 276 | - | polyphosphate kinase | YP_001231396 | 87 | 3E-112 | *Geobacter uraniireducens* Rf4 |

(Table S1). The constructed plasmid, further referred to as pBBR-*phyX*, was transformed into *E. coli* TOP10 cells (Invitrogen, UK). The newly generated strain was streaked on phytate medium to examine its phytate utilization phenotype. The pBBR-*phyX* indeed showed strong growth as compared with the pBBR1-MCS-4 vector control (Figs. 1A and 1B). This confirmed that the product of *phyX* is indeed involved in utilizing phytate as the sole P source. To confirm that the product of *phyX* possesses phytase enzymatic activity, an assay previously described by Yanke *et al.* [32] was used to monitor inorganic phosphate release from phytate by a cell-free extract of pBBR-*phyX* at 37°C. This showed that the cell-free extract derived from pBBR-*phyX* demonstrated a strong phytase activity based on the velocity of inorganic phosphate release per minute per gram of total protein (235 ± 40 μmol P/min/g). The pBBR1-MCS-4 vector control showed a background level of phytase activity (54 ± 16 μmol P/min/g) due to the *appA* gene of the *E. coli* host cells [6]. Data presented are the average of three independent biological replicates with standard error. This comparison showed that the phytase activity of pBBR-*phyX* was significantly higher than the background level, indicating that the PhyX product has indeed phytase activity. Interestingly, the optimum pH catalysis activity of pBBR-*phyX* was determined to be 6.5, whereas that of most HAP phytases is between 2.0 and 6.0 [31]. Since it has been reported that changing an amino acid closely adjacent to the RHGXRXP motif could make the optimum catalytic pH more neutral [5], a modification within the RHGXRXP motif could in fact have a similar influence. Compared with most typical HAP phytases, PhyX has a higher optimum pH, which could be more suitable for catalysis in slight acid and neutral environments, such as the gut of fish.

Within clone WFS_2, no ORF was identified to possess conserved motifs of any known phytase, suggesting that this clone possesses an unknown type of phytase. To locate the phytase gene, a subcloning strategy was adopted. Fosmid of WFS_2 was digested by *Hin*dIII and the restriction fragments were cloned into the *Hin*dIII site of pUC19 to generate a subcloning library. This library was screened on phytate medium to examine the phytate utilization capacity of the clones. One clone that exhibited phytate utilization capacity (Figs. 2A and 2B) was obtained. The recombinant plasmid, further referred to as pUC-SFPhy, was extracted from this clone and subjected to sequence analysis, which showed that the pUC19 vector was inserted with a subcloning fragment of 5.7 kb (further referred to as SFPhy). To confirm that the genes in the

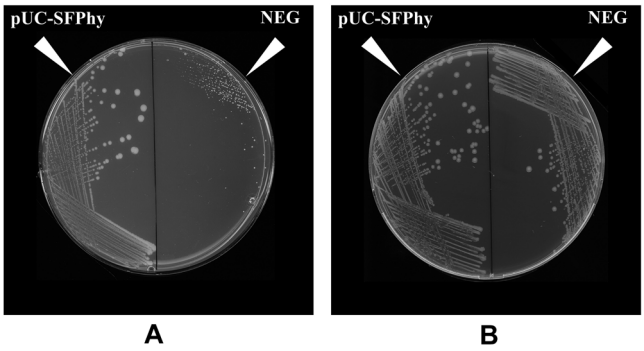


**Fig. 2.** Phytate utilization capacity of pUC-SFPhy in (**A**) medium supplemented with 5 mg/ml phytate, or (**B**) medium supplemented with 10 mM inorganic phosphate as the P source.
The pUC-SFPhy clone demonstrated a strong capacity to utilize phytate as a sole P source for growth, whereas the negative control clone (NEG, empty pUC19 vector) had slight growth due to an inevitable background phytase activity contributed by the product of the *E. coli appA* gene. Replacement of phytate with inorganic phosphate showed no growth differences.

SFPhy fragment are responsible for the phytase enzymatic activity, inorganic phosphate release from phytate by a cell-free extract was monitored, at its optimum pH of 7.0. This showed that the cell-free extract derived from pUC-SFPhy demonstrated a strong phytase activity (328 ± 32 μmol P/min/g). The pUC19 vector control showed a background phytase activity again (61 ± 16 μmol P/min/g, under the same conditions). Data presented are the average of three independent biological replicates with standard error. This comparison showed that pUC-SFPhy demonstrated a phytase activity significantly higher than the background level, indicating that the product of the ORFs in SFPhy has indeed phytase activity.

The nucleotide sequence of the SFPhy subcloning fragment was subjected to ORF prediction and annotation using RAST, which identified four ORFs (Table 2). Among these four ORFs, ORF_1 was annotated as a Crp/Fnr family transcription factor, while the other three (ORF_2, ORF_3, and ORF_4) were all annotated as hypothetical proteins showing very low similarity with known proteins. None of these four ORFs were related to known phytase families. This indicates that the ORFs in the SFPhy fragment might encode for a novel type of phytase. In order to identify the ORF responsible for the phytase activity of SFPhy, each of these four ORFs were individually subcloned into the pBBR1-MCS-4 expression vector (Table S1). The four constructed plasmids, further referred to as pBBR-SFPhy_ORF_1, pBBR-SFPhy_ORF_2, pBBR-SFPhy_ORF_3, and pBBR-SFPhy_ORF_4,

**Table 2.** Four ORFs in the restriction fragment subcloned into pUC-SFPhy, which were derived from fosmid WFS_2.

| ORF No. in pUC-SFPhy | Length of AA | Strand | Annotated function | Accession number of hit | Similarity of AA (%) | *E*-value | Taxonomic background of hit |
|---|---|---|---|---|---|---|---|
| 1 | 151 | - | Crp/Fnr family transcription factor | ACK43085 | 36 | 6E-19 | *Dictyoglomus turgidum* DSM 6724 |
| 2 | 101 | + | Hypothetical protein | EFI36146 | 34 | 2E-14 | *Desulfonatronospira thiodismutans* ASO3-1 |
| 3 | 404 | + | Hypothetical protein | ADY60128 | 37 | 4E-53 | *Planctomyces brasiliensis* DSM 5305 |
| 4 | 225 | + | Hypothetical protein | CAI06419 | 38 | 2E-27 | *Aromatoleum aromaticum* EbN1 |

The annotated function of each ORF (using RAST) and amino acid (AA) sequence similarity to its closest hits (by BLASTP) are shown.

were respectively transformed into *E. coli* TOP10 cells (Invitrogen). Interestingly, none of these four newly generated strains demonstrated phytase activity. This indicates that the product of each of the four ORFs could not work independently as a phytase. Instead, the phytase activity of SFPhy is more likely to require the participation of multiple ORFs. This also proved that the phytase activity demonstrated by SFPhy was not due to an up-regulation of the *appA* gene of *E. coli* by a transcription factor encoded by any of the four ORFs, particularly the ORF_1 that is likely to be a putative Crp/Fnr family transcription factor.

To test whether the putative transcription factor ORF_1 is involved in the expression of phytase activity of SFPhy, a mutant plasmid of pUC-SFPhy that removed the ORF_1 was constructed. For this, a modified fragment of SFPhy lacking 1,135 kb at the 5′ prime was cloned into the pUC19 vector (Table S1). This truncated SFPhy fragment was cloned in the same orientation and at the same restriction site in comparison with the original pUC-SFPhy. As a control, another mutant plasmid of pUC-SFPhy, which only lacked 111 bp at the 5′ prime of the SFPhy fragment, whereas ORF_1 remained intact, was also constructed using the same strategy (Table S1). These two plasmids, further referred to as pUC-SFPhy_DelF1135 and pUC-SFPhy_DelF111, were respectively transformed into *E. coli* TOP10 cells. Interestingly, pUC-SFPhy_DelF1135 lost the phytase activity, whereas UC-SFPhy_DelF111 possessing the intact ORF_1 retained the phytase activity. Since the ORF_1 seems not to be a transcription factor to up-regulate the *appA* of *E. coli*, these results altogether suggested that the ORF_1 might be a transcription factor that is essential for the multiple ORFs of SFPhy to demonstrate the phytase activity.

To further verify whether all four ORFs of SFPhy were essential for phytase activity, a random gene knockout strategy utilizing Tn5 *in vitro* transposition was used, as previously described by Martinez *et al.* [20]. Indeed, the result showed that all four ORFs were essential for pUC-SFPhy to retain its phytase activity. Thus, this putative novel phytase might either be a multi-subunit enzyme encoded by different structural genes, or consist of multiple enzymes that catalyze a multistep pathway of phytate degradation. This is different to known phytases, which are all encoded by a single structural gene [12, 17, 22]. The composition and structure of this putative novel phytase deserve further characterization.

This study revealed two novel type of phytases. They are very distinct in comparison with most known members of HAP, BPP, PAP, and CPhy phytases. Future efforts will be directed at characterizing these activities at the purified protein level. However, this study has demonstrated the potential of exploiting metagenomic approaches for mining new enzymatic activities with potential for industrial application. The novel phytases identified have the potential for agricultural and industrial applications to catalyze phytate degradation under neutral pH condition.

The nucleotide sequences of the *phyX* gene and the SFPhy fragment have been deposited to the Genbank database under accession numbers KF705200 and KF709432, respectively.

## Acknowledgments

This research was supported in part by grants awarded by the Science Foundation of Ireland (07/IN.1/B948, 12/TIDA/B2411, 12/TIDA/B2405, 09/RFP/BMT 2350); the Department of Agriculture, Fisheries and Food (FIRM/RSF/CoFoRD; FIRM 08/RDC/629); the Environmental Protection Agency (EPA 2008-PhD/S-2), the Irish Research Council for Science, Engineering and Technology (PD/2011/2414; RS/2010/2413), the European Commission (FP7-PEOPLE-2013-ITN, 607786; OCEAN2012, 287589; FP7-KBBE-2012-6, 311975; FP7-KBBE-2012-6, CP-TP-312184; Marie Curie 256596); and the Marine Institute (Beaufort

award C2CRA 2007/082); Teagasc (Walsh Fellowship 2013) and the Health Research Board (HRA/2009/146).

## References


1. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, *et al*. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9:** 75.
2. Chu HM, Guo RT, Lin TW, Chou CC, Shr HL, Lai HL, *et al*. 2004. Structures of *Selenomonas ruminantium* phytase in complex with persulfated phytate: DSP phytase fold and mechanism for sequential substrate hydrolysis. *Structure* **12:** 2015-2024.
3. Conry MJ, Hogan JJ. 2001. Comparison of cereals grown under high (conventional) and low (reduced) input systems. Teagasc, Crops Research Centre, Oak Park, Carlow.
4. Dvoráková J. 1998. Phytase: sources, preparation and exploitation. *Folia Microbiol.* **43:** 323-338.
5. Fu D, Huang H, Meng K, Wang Y, Luo H, Yang P, *et al*. 2009. Improvement of *Yersinia frederiksenii* phytase performance by a single amino acid substitution. *Biotechnol. Bioeng.* **103:** 857-864.
6. Greiner R, Konietzny U, Jany K-D. 1993. Purification and characterization of two phytases from *Escherichia coli*. *Arch. Biochem. Biophys.* **303:** 107-113.
7. Guyetant S, Giraud M, L'Hours L, Derrien S, Rubini S, Lavenier D, Raimbault F. 2005. Cluster of reconfigurable nodes for scanning large genomic banks. *Parallel Comput.* **31:** 7396.
8. Ha NC, Oh BC, Shin S, Kim HJ, Oh TK, Kim YO, *et al*. 2000. Crystal structures of a novel, thermostable phytase in partially and fully calcium-loaded states. *Nat. Struct. Biol.* **7:** 147-153.
9. 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.
10. Huang H, Luo H, Wang Y, Fu D, Shao N, Yang P, *et al*. 2009. Novel low-temperature-active phytase from *Erwinia carotovora* var. *carotovota* ACCC 10276. *J. Microbiol. Biotechnol.* **19:** 1085-1091.
11. Huang H, Shi P, Wang Y, Luo H, Shao N, Wang G, *et al*. 2009. Diversity of beta-propeller phytase genes in the intestinal contents of grass carp provides insight into the release of major phosphorus from phytate in nature. *Appl. Environ. Microbiol.* **75:** 1508-1516.
12. Huang H, Zhang R, Fu D, Luo J, Li Z, Luo H, *et al*. 2011. Diversity, abundance and characterization of ruminal cysteine phytases suggest their important role in phytate degradation. *Environ. Microbiol.* **13:** 747-757.
13. Jorquera M, Martínez O, Maruyama F, Marschner P, de la Luz Mora M. 2008. Current and future biotechnological applications of bacterial phytases and phytase-producing bacteria. *Microb. Environ.* **23:** 182-191.
14. Knowlton KF, McKinney JM, Wilson KF, Cobb C. 2003. Effect of an exogenous phytase enzyme blend and dietary phosphorus content on P excretion in lactating cows. *J. Dairy Sci.* **86:** 224.
15. Konietzny U, Greiner R. 2004. Bacterial phytase: potential application, *in vivo* function and regulation of its synthesis. *Braz. J. Biol.* **35:** 11-18.
16. Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, Roop RM, Peterson KM. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166:** 175-176.
17. 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.
18. Lee DY, Schroeder J, Gordon DT. 1988. Enhancement of Cu bioavailability in the rat by phytic acid. *J. Nutr.* **118:** 712-717.
19. Lei X, Stahl CH. 2001. Biotechnological development of effective phytases for mineral nutrition and environmental protection. *Appl. Microbiol. Biotechnol.* **57:** 474-481.
20. Martínez A, Bradley AS, Waldbauer JR, Summons RE, Delong EF. 2007. Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host. *Proc. Natl. Acad. Sci. USA* **104:** 5590-5595.
21. Mukhametzyanova AD, Akhmetova AI, Sharipova MR. 2012. Microorganisms as phytase producers. *Microbiology* **81:** 267-275.
22. Mullaney EJ, Ullah AH. 2003. The term phytase comprises several different classes of enzymes. *Biochem. Biophys. Res. Commun.* **312:** 179-184.
23. Patel KJ, Vig S, Kumar GN, Archana G. 2010. Effect of transgenic rhizobacteria over-expressing *Citrobacter braakii appA* on phytate–P availability to mung bean plants. *J. Microbiol. Biotechnol.* **20:** 1491-1499.
24. Selle PH, Cowieson AJ, Ravindran V. 2009. Consequences of calcium interactions with phytate and phytase for poultry and pigs. *Livestock Sci.* **124:** 126-141.
25. Selle PH, Ravindran V. 2008. Phytate-degrading enzymes in pig nutrition. *Livestock Sci.* **113:** 99-122.
26. Seo M-J, Kim J-N, Cho E-A, Park H, Choi H-J, Pyun Y-R. 2005. Purification and characterization of a novel extracellular alkaline phytase from *Aeromonas* sp. *J. Microbiol. Biotechnol.* **15:** 745-748.
27. Shin S, Ha NC, Oh BC, Oh TK, Oh BH. 2001. Enzyme mechanism and catalytic property of β propeller phytase. *Structure* **9:** 851-858.
28. Thacker PA, Rossnagel BG, Raboy V. 2004. Effect of phytase supplementation on phosphorus digestibility in low-phytate barley fed to finishing pigs. *Arch. Anim. Nutr.* **58:** 61-68.

29. Unno Y, Okubo K, Wasaki J, Shinano T, Osaki M. 2005. Plant growth promotion abilities and microscale bacterial dynamics in the rhizosphere of Lupin analysed by phytate utilization ability. *Environ. Microbiol.* **7:** 396-404.
30. Van Etten RL, Davidson R, Stevis PE, MacArthur H, Moore DL. 1991. Covalent structure, disulfide bonding, and identification of reactive surface and active site residues of human prostatic acid phosphatase. *J. Biol. Chem.* **266:** 2313-2319.
31. Vats P, Banerjee UC. 2004. Production studies and catalytic properties of phytases: an overview. *Enzyme Microb. Technol.* **35:** 3-14.
32. Yanke LJ, Bae HD, Selinger LB, Cheng KJ. 1998. Phytase activity of anaerobic ruminal bacteria. *Microbiology* **144:** 1565-1573.
33. Yao B, Shao N, Huang H, Meng K, Luo H, Wang Y, Yang P. 2008. Cloning, expression, and characterization of a new phytase from the phytopathogenic bacterium *Pectobacterium wasabiae* DSMZ 18074. *J. Microbiol. Biotechnol.* **18:** 1221-1226.
34. Yao MZ, Zhang YH, Lu WL, Hu MQ, Wang W, Liang AH. 2012. Phytases: crystal structures, protein engineering and potential biotechnological applications. *J. Appl. Microbiol.* **112:** 1-14.