Vol.58, n.3: pp. 337-342, May-June 2015
http://dx.doi.org/10.1590/S1516-8913201500040
ISSN 1516-8913 Printed in Brazil

BRAZILIAN ARCHIVES OF
BIOLOGY AND TECHNOLOGY
A N I N T E R N A T I O N A L J O U R N A L

Protease Gene Shuffling and Expression in *Pichia pastoris*

Gang Yang[1], Juan Chang[1], Qingqiang Yin[1*], Erzhu Wang[1], Qun Zhu[1], Ping Wang[1], Xiaowei Dang[2] and Fushan Lu[3]

[1]College of Animal Science and Veterinary Medicine; Henan Agricultural University; Zhengzhou - China. [2]Henan Delin Biological Product Co. Ltd.; Xinxiang - China. [3]Henan Engineering and Technology Research Center of Feed Microbes; Zhoukou - China

ABSTRACT

Four kinds of neutral and alkaline protease genes from Aspergillus oryzae *and* Bacillus subtilis *were isolated and shuffled. The shuffled genes were selected, inserted into pGAPZαA plasmid and transformed into* Escherichia coli. *The gene which could express high-activity protease was selected by screening the sizes of transparent zones around the colonies on casein plates. After an ideal protease gene was selected, it was sequenced and then transformed into* Pichia pastoris *X33. The result showed that the base in 1022th position of shuffled protease gene was changed from thymine to cytosine, inferring that cysteine was changed to arginine in the mutant protease. After 48 h incubation for the transformed* P. pastoris *with the mutant or native protease genes, the mutant protease activity was 36.4% higher than the native protease (P<0.05). The optimal pH and temperature of the mutant protease were 6.5-8.0 and 30-70°C, respectively, which indicated better stability than the native protease (P<0.05).*

Key words: Protease gene, DNA shuffling, *Pichia pastoris*, gene expression, protease characterization

INTRODUCTION

Protease has a big market in the world with an estimated market of about three billion US dollars annually (Leary et al. 2009) due to its important roles in improving nutrient values, retarding deterioration, preventing undesired interactions, decreasing flavors and odors, and removing toxic or inhibitory factors (Pardo et al. 2000). In the past decades, genetic engineering methods have been used in enzyme industry to improve protease activity and characterization (Farag and Hassan 2004; Guo and Ma 2008; Zhang et al. 2014).

DNA shuffling has been used to recombine homologous DNA sequences during molecular evolution *in vitro* since 1994 (Stemmer 1994). This technique has been shown to be useful in combining independently isolated mutations of a gene into a single progeny. It has also been widely applied in scientific studies and the recombination of many types of proteins such as gene medicine and enzymes (Li et al. 2007; Uesugi et al. 2011; Akbulut et al. 2013). The directed evolution involves the random mutagenesis of one or more starting enzymatic genes, followed by a screening or selection step to isolate or enrich enzyme variants with improvements in one or more desirable properties. Therefore, an efficient high-throughput screening method is a critical step for the success of directed evolution studies.

Attempts have continued to improve the production and characteristics of proteases to meet the demand. Generally, protease gene plays an important role in protease production and characterization. Even though protease gene expressions have been done (Yang et al. 2013; Banani et al. 2014), the protease gene mutation by gene shuffling has not been studied. The purpose

*Author for correspondence: qqy1964@126.com

of this work was to develop an ideal mutant protease gene and characterize the recombinant enzyme activities.

MATERIAL AND METHODS

Microbial strains, *plasmids* and incubating conditions

Aspergillus oryzae CGMCC5817 and *Bacillus subtilis* ACCC10619 were preserved in the laboratory. *Escherichia coli*, DNA polymerase, pMD19-T vector, all restriction endonucleases and T4 ligase were purchased from Takara (Dalian, China). *Pichia pastoris* X33, pGAPZαA and zeocin were purchased from Invitrogen, USA.

Aspergillus oryzae was cultivated in a medium containing (%) 0.6 soluble starch, 2.0 glucose, 0.2 yeast extract, 0.5 peptone, 0.2 KH_2PO_4 and 0.03 $MgSO_4$ by incubating at 30°C and 120 rpm for 72h. *B. subtilis* was grown in LB liquid medium (1% peptone, 0.5% yeast extract, 1.0% NaC1) by incubating at 37°C and 120 rpm for 48h. *E. coli* was grown in LB or low-salt LB medium (0.5% NaC1) by incubating at 37°C and 120 rpm for 48h. *Pichia pastoris* was grown in YPD medium (2% glucose, 2% peptone, 1% yeast extract) by incubating at 30°C and 120 rpm for 48h. Agar (2%) and antibiotics were added in the above media when necessary. All the media were autoclaved at 121°C for 20 min.

Isolation of four protease genes from *A. oryzae* and *B. subtilis*

The cells of *A. oryzae* and *B. subtilis* were collected by centrifuge, frozen in liquid nitrogen, and then ground in mortar. The total RNA was extracted by using RNA Extract Kit (Takara, Dalian, China). Double strand cDNA of four protease genes were synthesized by RT-PCR. The First Strand cDNA Synthesis Kit (TransGen, Beijing, China) was used in reverse transcription. This transcription reaction mixture contained 2.0 μL total RNA, 10 μL 2×ES reaction mix, 1.0 μL Oligo$(dT)_{18}$, 1.0 μL RT enzyme mix and 6.0 μL RNase-free water, which was carried out at 42°C for 30 min and 99°C for 5 min.

Four kinds of protease genes in GenBank are listed in Table 1. The primers in Table 2 were designed and synthesized based on the nucleotide sequence of the above four genes. PCR was carried out in a total volume of 50 μL mixture containing 2.0 μL reverse transcription product as template, 25 μL PCR SuperMix, 1.0 μL up-stream primer, 1.0 μL down-stream primer, 21 μL ddH_2O. The PCR process was initiated at 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 62°C for 30 s, 72°C for 100 s, and a final extension at 72°C for 10 min. The PCR products were analyzed by electrophoresis with 1.0% agarose gels stained with ethidium bromide and purified by DNA gel recovery Kit (TransGen, Beijing, China).

Table 1 - Protease genes from *Aspergillus oryzae and B. subtilis*.

Genes	GenBank No.	DNA length
Neutral protease gene from *A. oryzae* (NPGA)	S53810.1	1059 bp
Alkaline protease gene from *A. oryzae* (APGA)	S75278.1	1212 bp
Neutral protease gene from *B. subtilis* (NPGB)	AB568089.1	1566 bp
Alkaline protease gene from *B. subtilis* (APGB)	K01985.1	1149 bp

Table 2 - PCR Primers of four protease genes.

Genes	Primers	DNA sequences	Endonucleases
NPGA	Up stream	CCGGAATTCATGCGTGTCACTACTCTCTCC	*EcoR I*
	Down stream	GCTCTAGATTAGCACTTGAGCTCGATAGC	*Xba I*
APGA	Up stream	TGCTCTAGAATGCAGTCCATCAAGCGTACC	*Xba I*
	Down stream	CCGCTCGAG TTAAGCGTTACCGTTGTAGGC	*Xho I*
NPGB	Up stream	GCTCTAGAGTGGGTTTAGGTAAGAAATTGTCTG	*Xba I*
	Down stream	CCGCTCGAGTTACAATCCAACAGCATTC	*Xho I*
APGB	Up stream	CCGCTCGAGGTGAGAGGCAAAAAGGTATGG	*Xho I*
	Down stream	GCTCTAGATTACTGAGCTGCCGCCTGTAC	*Xba I*

DNA Shuffling

The purified PCR products was digested in 50 μL mixture containing 10 μL neutral protease gene from *A. oryzae* (NPGA), 10 μL alkaline protease gene from *A. oryzae* (APGA), 10 μL neutral protease gene from *B. subtilis* (NPGB), 10 μL alkaline protease from *B. subtilis* (APGB), 5.0 μL 10×Dnase I buffer and 5.0 μL DNase I (0.02

U/μL) for 10 min at 15°C and 10 min at 80°C. The fragments of about 50-base pairs were purified by DNA gel recovery kit with 2% agarose gel.

The PCR reaction without primers containing 20 μL gene fragment and 20 μL PCR SuperMix was initiated at 94°C for 5 min, followed by 45 cycles of 94°C for 30 s, 50°C for 30 s, 72°C for 30 s, and a final extension at 72°C for 10 min. After reaction, 5.0 μL PCR product was used as template and added into 45 μL PCR mixture containing 1.0 μL NPGA up-stream primer, 1.0 μL NPGA down-stream primer, 25 μL PCR SuperMix and 18 μL ddH_2O. The PCR reaction was 94°C for 5 min, and then 35 cycles of 94°C for 30 s, 62°C for 30 s, 72°C for 80 s, and finally 72°C for 10 min. The PCR product was analyzed by electrophoresis with 1.0% agarose gel, and purified by DNA gel recovery kit.

Protease gene transformation in *E. coli* and target gene screening

The purified PCR product was ligated with pMD19-T vector and transformed into *E. coli* competent cells by heat shock method, and the transformants were selected on LB plates containing 0.4% casein, ampicillin (100 mg/mL), IPTG (24 mg/mL) and X-gal (20 mg/mL). The transparent zone was used as an indicator of protease activity. The colonies with the largest transparent zones were selected and inoculated in LB medium at 37°C for 48 h. The vector with mutant protease gene was extracted by plasmid recovery kit. DNA sequence was determined in Sangon (Shanghai, China) and analyzed with DNAMAN (Lynnon Biosoft, USA).

Protease gene expression in *P. pastoris*

The vector with mutant protease gene was digested by double endonucleases (E*coR* I and X*ba* I). The protease gene was isolated with 1% agarose gel, purified with kit, and then ligated with pGAPZαA expression vector pre-digested with the same endonucleases to construct pGAPZaA-mutant protease gene vector. After the vector was transformed into *E. coli* 5Hα by heat shock method, the transformants were selected on low salt LB plates containing zeocin (100 μg/mL). The vector was isolated and purified with kit, linearized with endonuclease (Bln I), and then transformed into competent *P. pastoris* with electroporation. The competent cells were prepared as follows: One fresh colony was selected and incubated overnight in 5.0 mL of YPD medium in a 50 mL flask at 30°C, and then 0.1 mL was taken and put into 100 mL fresh YPD medium in a 500 mL flask for incubating until an optical density of 1.0 at 600 nm. The cells were centrifuged (1500 g) at 4°C for 5 min, suspended two times with ice-cold sterile water and two times with ice-cold 1 M sorbitol. Finally, the cells were re-suspended in 1.0 mL ice-cold 1 M sorbitol. Eighty μl cell suspensions were mixed with 5-10 μg linearized pGAPZaA-protease gene and transferred into an ice-cold 0.2 cm electroporation cuvette. Parameters used for electroporation were 1.5 kV/cm, 25 μF and 400Ω. The transformed *P. pastoris* was sprayed on YPD plates with zeocin (100 μg/mL), incubated at 30°C for 48 h, and the positive colonies were selected for the next step. The pGAPZαA with the native protease gene was also transformed into *P. pastoris* as control.

Protease activity and characterization analysis

The recombined *P. pastoris* was incubated overnight in YPD medium containing zeocin (100 μg/mL). About 0.4% recombinant *P. pastoris* incubation was added into new YPD medium and shaken at 200 rpm for 24, 48 and 72 h. The medium was centrifuged at 8000 g for 10 min, and the supernatant was used for protease activity analysis with Folin method (Oda and Murao 1974). One unit of enzyme activity was defined as liberating 1 μg/mL tyrosine per min.

The native and mutant protease activities at different pH points and temperatures were determined. To determine the optimal pH value of protease, NaH_2PO_4-Na_2HPO_4 buffers (pH 6.0, 6.5, 7.0, 7.8 and 8.0) were used. The optimal temperature was determined based on the temperature ranges of 30, 40, 50, 60 and 70°C at pH 7.0 for 10 min.

Data analysis

The data were analyzed using the ANOVA procedures of Statistical Analysis Systems institute (SAS 8.0). Duncan's multiple range test was used to evaluate treatment means. The results were considered statistically significance at $P<0.05$.

RESULTS

Amplification and homological analysis of four protease genes from *A. oryzae* and *B. subtilis*

Four genes encoding neutral and alkaline proteases of *A. oryzae* and *B. subtilis* were amplified by RT-PCR, respectively (Fig.1). The homology of four

genes was 42.69% analyzed by DNAMAN software.

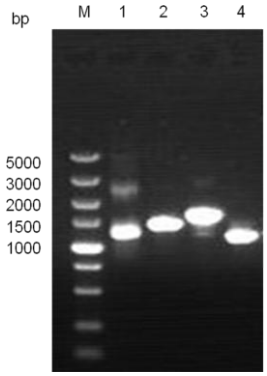

Figure 1 - PCR products of four protease genes. Lane M: standard DNA marker, Lane 1: NPGA, Lane 2: APGA, Lane 3: NPGB, Lane 4: APGB.

DNA shuffling of protease genes

Four genes were mixed equally and digested into 50-100 bp fragments by Dnase I (Fig. 2). The small gene fragments was linked up to 300-500 bp by PCR without primers, and then amplified by PCR with the primers of NPGA. The PCR product of mutant protease gene was analyzed by 1.0% agarose gel (Fig. 3).

The ideal mutant protease gene selection and analysis

After protease gene transformation in *E. coli*, the positive transformants were selected by blue-white assay and casein plate. The best transformant was selected, and the ideal mutant protease gene was sequenced. The result showed that the base in 1022th position of mutant protease gene was changed from thymine to cytosine, inferring that cysteine was changed to arginine in the mutant protease.

PCR identification of protease gene in the recombinant *P. pastoris*

Three kinds of genomic DNA were isolated as templates from three kinds of recombinant *P. pastoris* with native and mutant protease genes or without protease gene, respectively. The PCR products of three kinds of recombinants were analyzed by 1.0% agarose gel (Fig. 4). The results indicated that the native and mutant protease genes were expressed in *P. pastoris* successfully.

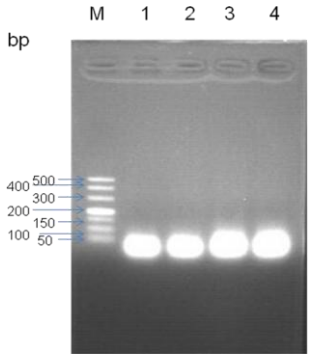

Figure 2 - Small DNA fragment analysis of four protease genes. Lane M: standard DNA marker; Lane 1-4: NPGA, APGA, NPGB and APGB gene fragments digested by Dnase I.

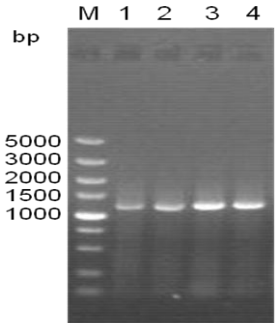

Figure 3 - PCR product with the primers of NPGA gene. Lane M: standard DNA marker; Lane 1-4: PCR products.

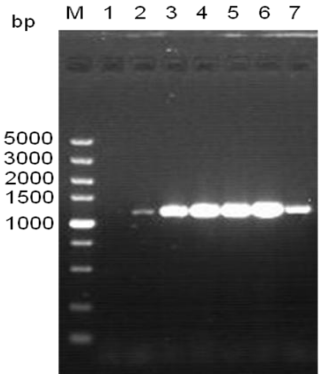

Figure 4 - The protease gene identification of recombinant *P. pastoris* by PCR. Lane M: standard DNA marker; Lane 1: the recombinant *P. pastoris* without protease gene; Lane 2: the recombinant *P. pastoris* with native NPGA gene; Lane 3-6: the recombinant *P. pastoris* with mutant protease gene; Lane 7: the native NPGA gene.

Protease activity measurement and characterization analysis

Protease activities in the supernatant of the recombinant *P. pastoris* was listed in Table 3. After 48 h incubation for the transformed *P. pastoris* with the mutant or native protease genes, the mutant protease activity was 36.4% higher than the native protease (P<0.05). The optimal pH values of the mutant protease were 6.5-8.0, which had larger optimal pH ranges than the native protease (pH 7.0-8.0) for keeping relative protease activity more than 80% (*P*<0.05, Fig. 5). The optimal temperature of the mutant protease was 30-70°C, which had larger optimal temperature ranges than the native protease (50°C) for keeping relative protease activity more than 80% (*P*<0.05, Fig. 6). It was indicated that the mutant protease was more thermo-stable than the native one.

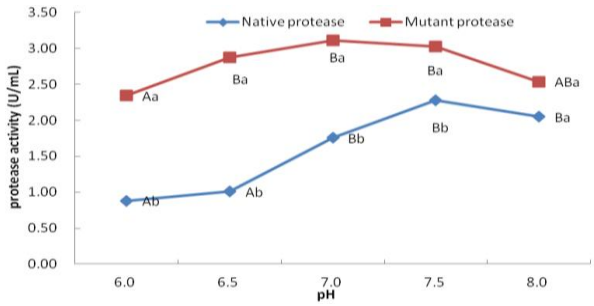

Figure 5 - The optimal pH values of native and mutant proteases. The different capital letters mean significant differences between the different pH points for each protease, respectively (*P*<0.05); while the same capital letters mean insignificant differences (*P*>0.05). The different lower-case letters mean significant differences between two proteases at the same pH points (*P*<0.05), while the same lower-case letters mean insignificant differences (*P*>0.05).

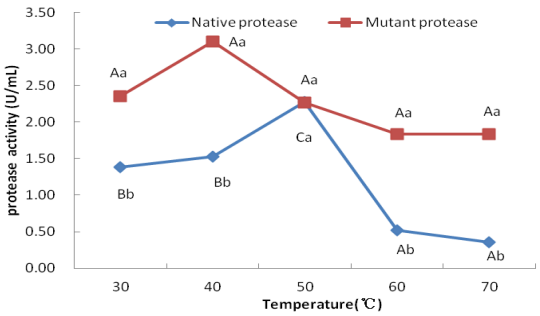

Figure 6 - The optimal temperature of native and mutant proteases. The different capital letters mean significant differences between the different temperature points for each protease, respectively (*P*<0.05); while the same capital letters mean insignificant differences (*P*>0.05). The different lower-case letters mean significant differences between two proteases at the same temperature points (*P*<0.05), while the same lower-case letters mean insignificant differences (*P*>0.05).

Table 3 - Protease activity of the recombinants during different periods of incubation (U/mL).

Incubation time (h)	24	48	72
Recombinant *P. pastoris* with native protease gene	1.22±0.20A	2.28±0.19B	0.75±0.15B
Recombinant *P. pastoris* with mutant protease gene	1.32±0.16A	3.11±0.11A	1.30±0.15A

Each value represents mean ± SE of five replicates per treatment. The different capital letters in the same columns mean significant difference (*P*<0.05), while the same capital letters in the same columns mean insignificant difference (*P*>0.05).

DISCUSSION

Since DNA shuffling technique was introduced in 1994, it has been improved from single gene shuffling to family gene shuffling (Stemmer 1994). Family gene shuffling utilizes naturally occurring nucleotide substitutions among family genes as the driving force for evolution in vitro, which can improve the efficiency of DNA shuffling as well as gene mutation probability. The application of family gene shuffling strategy has been demonstrated by the previous reports (Crameri et al. 1998; Niederhauser et al. 2012; Madan and Mishra 2014). It was indicated that the gene sequence homology of parent template should be more than 70% for gene shuffling in order to achieve the favorable result (Joern et al. 2002). Although the sequence homology of four protease genes in this study was only 42%, the good result was achieved by increasing PCR cycles from 15 to 40 without primers, indicating that the higher homology was not essential for DNA shuffling.

Compared with the native protease, the activity, pH ranges and thermostability of mutant protease was improved significantly, probably due to the modification of protease gene structure. It has been proposed that several sequences or structural features contribute to the greater stability and activity of proteins. These features include packing of the core structure (Kumar et al. 2000), hydrogen bond (Zuo et al. 2007), surface hydrophobic

interactions (Niu et al. 2006), salt bridge (Xie et al. 2006), and so on. After DNA shuffling in this study, cysteine was mutated to arginine in neutral protease. It has been reported that arginine could contribute the addition of hydrogen bonds, which presumably participate in the ground state binding of substrate and contribute to transition state stabilization for improving the specific activity of protease (Guo et al. 2014). In addition, because the mutation position (341^{th} site) was close to active site of protease (304^{th} site, Mcauley et al. 2001), the hydration layer formed by hydrogen bonds and big steric bulk of arginine could improve pH ranges and thermostability.

CONCLUSION

The mutant protease gene was obtained by shuffling four protease genes from two species of microbes. After the mutant gene was cloned and expressed in *P. pastoris*, the mutant protease indicated larger ranges of pH, higher activity and thermostability than the native protease. The ideal characterization of mutant protease could be useful for commercial application.

REFERENCES

Akbulut N, Öztürk MT, Pijning T, Öztürka S, Gümüsel F. Improved activity and thermostability of *Bacillus pumilus* lipase by directed evolution. *J Biotechnol*. 2013; 164: 123-129.

Banani H, Spadaro D, Zhang D, Matic S, Garibaldi A, Gullino ML. Biocontrol activity of an alkaline serine protease from *Aureobasidium pullulans* expressed in *Pichia pastoris* against four postharvest pathogens on apple. *Int J Food Microbiol*. 2014; 182-183: 1-8.

Crameri A, Raillard SA, Bermudez E, Stemmer WPC. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature*. 1998; 391: 288-291.

Farag AM, Hassan MA. Purification, characterization and immobilization of a keratinase from *Aspergillus oryzae*. *Enzyme Microb Tech*. 2004; 34: 85-93.

Guo FF, Zhang C, Bie XM, Zhao HZ, Diao HW, Lu FX, et al. Improving the thermostability and activity of lipoxygenase from *Anabaena* sp. PCC 7120 by directed evolution and site-directed mutagenesis. *J Mol Catal B-Enzym*. 2014; 107: 23-30.

Guo JP, Ma Y. High-level expression, purification and characterization of recombinant *Aspergillus oryzae* alkaline protease in *Pichia pastoris*. *Protein Expres Purif*. 2008; 58: 301-308.

Joern JM, Meinhold P, Arnold FH. Analysis of shuffled gene libraries. *J Mol Biol*. 2002; 316: 643-656.

Kumar S, Tsai CJ, Nussinov R. Factors enhancing protein thermostability. *Protein Eng*. 2000; 13: 179-191.

Leary D, Vierros M, Hamon G, Arico S, Monagle C. Marine genetic resources: a review of scientific and commercial interest. *Mar Policy*. 2009; 33: 183-194.

Li J, Shi PJ, Han XY, Meng K, Yang PL, Wang YR, et al. Functional expression of the keratinolytic serine protease gene sfp2 from *Streptomyces fradiae var.* k11 in *Pichia pastoris*. *Protein Expres Purif*. 2007; 54: 79-86.

Madan B, Mishra P. Directed evolution of *Bacillus licheniformis* lipase for improvement of thermostability. *Biochem Engin J*. 2014; 91: 276-282.

Mcauley KE, Jiaxing Y, Dodson EJ, Lehmbeck J, Østergaard PR, Wilson KS. A quick solution: ab initio structure determination of a 19 kDa metalloproteinase using ACORN. *Acta Crystallogr D Biol Crystallogr*. 2001; 57: 1571-1578.

Niederhauser B, Siivonen J, Määttä JA, Jänis J, Kulomaa MS, Hytönen VP. DNA family shuffling within the chicken avidin protein family–A shortcut to more powerful protein tools. *J Biotechnol*. 2012; 157: 38-49.

Niu WN, Li ZP, Tan TW. Secretion of pro- and mature *Rhizopus arrhizus* lipases by *Pichia pastoris* and properties of the proteins. *Mol Biotechnol*. 2006; 32: 73-81.

Oda K, Murao S. Purification and some properties of acid proteinase A and B of *Scytalidium lignicolum* ATCC 24568. *Agr Biol Chem*. 1974; 38: 2435-2444.

Pardo MF, Lopez MI, Canals F, Zviles FX, Natalucci CL, Caffini NO. Purification of balansain I, an endopeptidase from unripe fruits of *Bromelia balansae Mez* (*Bromeliaceae*). *J Agric Food Chem*. 2000; 48: 3795-3800.

Stemmer W. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature*. 1994; 370: 389-391.

Uesugi Y, Usuki H, Arima J, Iwabuchi M, Hatanaka T. Molecular dissection of *Streptomyces* trypsin on substrate recognition. *Biochim Biophys Acta*. 2011; 1814: 1295-1304.

Xie HF, Flint J, Vardakou M, Lakey JH, Lewis RJ, Gilbert HJ, et al. Probing the structural basis for the difference in thermostability displayed by family 10 xylanases. *J Mol Biol*. 2006; 360: 157-167.

Yang X, Cong H, Song J, Zhang J. Heterologous expression of an aspartic protease gene from biocontrol fungus *Trichoderma asperellum* in *Pichia pastoris*. *World J Microbiol Biotechnol*. 2013; 29: 2087-2094.

Zhang H, Zhang B, Zheng YB, Shan AS, Cheng BJ. Neutral protease expression and optimized conditions for the degradation of blood cells using recombinant *Pichia pastoris*. *Int Biodeter Biodegra*. 2014; 93: 235-240.

Zuo ZY, Zheng ZL, Liu ZG, Yi QM, Zou GL. Cloning, DNA shuffling and expression of serine hydroxymethyltransferase gene from *Escherichia coli* strain AB90054. *Enzyme Microb Tech*. 2007; 40: 569-577.

Received: January 31, 2015;
Accepted: March 30, 2015.