Int J Clin Exp Med 2018;11(11):11804-11815
www.ijcem.com /ISSN:1940-5901/IJCEM0072609

Original Article

Cathepsin D as a new allergen from *Dermatophagoides farinae*

Jianli Lin[1,2*], Chengshen Hu[3*], Junxian He[3*], Qinghui Fu[3], Nana Huang[3], Desheng Chen[3], Wingfu Lai[4], Xiaoyu Liu[2], Zhigang Liu[1,2]

[1]State Key Laboratory of Respiratory Disease, National Clinical Center for Respiratory Diseases, Guangzhou Institute of Respiratory Diseases, First Affiliated Hospital, Guangzhou Medical University, Guangzhou, Guangdong Province, China; [2]Department of Allergy, Affliated Luohu Hospital of Shenzhen University, Shenzhen Luohu Hospital Group, Shenzhen, China; [3]State Key Laboratory of Respiratory Disease for Allergy at Shenzhen University, Shenzhen Key Laboratory of Allergy and Immunology, School of Medicine, Health Science Center, Shenzhen University, Shenzhen, China; [4]School of Pharmaceutical Science, Health Science Center, Shenzhen University, Shenzhen, China. [*]Equal contributors.

Received September 11, 2017; Accepted June 20, 2018; Epub November 15, 2018; Published November 30, 2018

Abstract: House dust mites, *Dermatophagoides farinae* (*D. farina*), are a major indoor allergen causing allergies. Allergic diseases caused by house dust mites have become a worldwide problem. Proper characterization of allergens is pivotal to the development of effective diagnosis or immunotherapy. This present study reports a new allergen, cathepsin D, extracted from *D. farina.* To evaluate the allergic activity of cathepsin D, total RNA was extracted and cathepsin D was amplified by RT-PCR. Results of restriction enzyme digestion indicated that the cathepsin D gene consisting of 1,143 bases was cloned. SDS-PAGE showed that the size of cathepsin D was 43 kDa. After expression in *E. coli* and affinity purification, IgE reactivity was evaluated by enzyme-linked immune sorbent assay (ELISA), with the sera of mite-related allergic patients, and by skin prick tests (SPT) in mite-related allergic human subjects. Positive results in ELISA and SPT revealed that cathepsin D has allergenicity and is a new allergen found in *D. farina.*

Keywords: *Dermatophagoides farina*, cathepsin D, allergen, purification, allergenicity

Introduction

Prevalence of respiratory allergic diseases has increased rapidly over the past several decades [1]. House dust mites (HDMs) are commonly found in the environment. Allergens from house dust mites have been recognized as a major cause of airway allergic diseases, accounting for more than 50% of asthma cases and/or rhinitis in patients sensitized to house dust mites [2]. Specific immunotherapy (SIT) is considered to be the only disease-modifying treatment for allergic diseases [3]. At present, immunotherapeutic vaccines for allergies are largely prepared from crude allergen extracts. However, the use of crude extracts may lead to numerous complications. One of these complications is caused by the fact that crude extracts often contain ill-defined allergens and could be polluted by nonallergenic components. This presents a technical barrier to the development of SIT [4]. Thirty-three groups of dust mite allergens have been identified from *Dermatophagoides farinae* (http://www.allergen.org/). Some of these have been extensively studied, including Der f1 and Der f2. Apart from these two groups of allergens, Der f3, f6, f7, and other allergens may provoke hyperreactivity or inflammatory reactions [5-11, 12]. Therefore, identification of allergens in *D. farina* is essential to improving efficacy and safety of SIT with HDMs.

Previous studies have shown that cathepsin D is an aspartic protease present in lysosome-related organelles, such as secretory inflammatory granules of basophilic/mast cells, secretory lytic granules of natural killer cells, and endosome-like MHC compartment of antigen-presenting cells (APCs) [13]. Cathepsin D is involved in many pathological processes, including inflammatory states, neoplasms, rheumatic diseases, apoptosis, and muscle dystrophy. It has also been reported to be closely

associated with allergic asthma [14, 15]. However, the allergenicity of cathepsin D from *Dermatophagoides farinae* remains unclear. This present study cloned, expressed, and purified cathepsin D. SPT tests were also used to evaluate levels of allergenicity of cathepsin D. Results of this study demonstrate that cathepsin D is capable of inducing hypersensitivity in humans.

The concept of precision medicine is currently emerging [16, 17]. This emergence is strongly promoted by the development of bioinformatics. This is because the use of bioinformatics can help predict key target spots of a disease, providing important references for drug development. Previous reports about epitope peptide vaccines have shown excellent potential application prospects, hence, epitope prediction of novel allergens is required [18-20]. Novel vaccines, synthesized based on B-cells of the T-cell epitope, have been demonstrated to activate human B cells to produce specific antibodies, activating cytotoxic lymphocytes (CTL) to eliminate infected cells [18]. In this study, diverse bioinformatics tools were used to predict properties of cathepsin D, providing valuable information for further vaccine development and clinical applications.

Materials and methods

Sera of allergic patients

Sera of allergic patients were supplied from the First Affiliated Hospital of Guangzhou Medical University. This study was approved by the Human Ethics Committee of Shenzhen University.

Culture and extraction of house dust mites

House dust mites were cultured at 25°C with 70% relative humidity, as previously reported [21]. After several days of culturing, ripe mites were extracted using a modified heat-escape method. After washing with 1X PBS, the bodies of mites were grinded to powder in liquid nitrogen. After treatment with lysate (9M urea, 4% CHAPS, 60 mM DTT, 2% IPG buffer), samples were centrifuged at 15,000 rpm for 50 minutes at 4°C. The supernatant was stored after lyophilization.

Total RNA extraction and cDNA cloning of cathepsin D

Total RNA was extracted from treated mite bodies with TRIzol (Invitrogen), according to manufacturer instructions. To synthesize full-length cDNA of cathepsin D, reverse transcription-polymerase chain reaction assay (RT-PCR) was used with extracted total RNA as a template. Each reaction system of 50 μL consisted of 5 μL 10 x ExTaq Buffer, 0.25 μL TaKaRa ExTaq, and a mixture of 4 μL dNTP and 2 μL primers of 5' and 3'. Remaining parts of the solution were supplemented with deionized water. PCR program was set with the following conditions: denaturation at 94°C for 1 minute, annealing at 50°C for 1 minute, and extension at 72°C for 1 minute, for 35 cycles. Completion of target cDNA cloning would be ensured by the Bioanalyzer and 1% agarose electrophoresis.

Expression and purification of recombinant cathepsin D

cDNA purified products were connected to a PMD-18T vector and the constructed plasmid was transformed into *E. coli* by heat transformation. Transformed bacteria were cultured in the LB-agarose medium with 100 mg/L ampicillin overnight at 37°C. Afterward, recombinant *E. coli* were collected for plasmid extraction. Extracted plasmids were assayed by BamH I enzymes and were delegated to BGI (Shenzhen, China) for sequencing. The target sequence corresponding to cathepsin D was ligated with the pET-28a (+) vector at 37°C for 4 hours. Next, recombinant cloning products were amplified with *E. coli* using the same method as mentioned above. To express the target protein, Pet-28a (+)-cathepsin D plasmid was translated to BL21 *E. coli* and was then screened in the kanamycin LB-agarose medium. The *E. coli* were then cultured in the LB-medium supplemented with 50 μg/mL kanamycin until the logarithmic phase (A600 nm = 0.6~0.9) was reached. Isopropyl-D-thiogalactopyranoside (IPTG) was supplemented to the culture medium to induce expression of the target protein. Bacterial cells were cultured at 37°C for 4 hours and harvested by centrifugation. After ultrasonic fragmentation, the supernatant and precipitate of the solution were analyzed with sodium dodecyl sulfate-agarose gel elec-

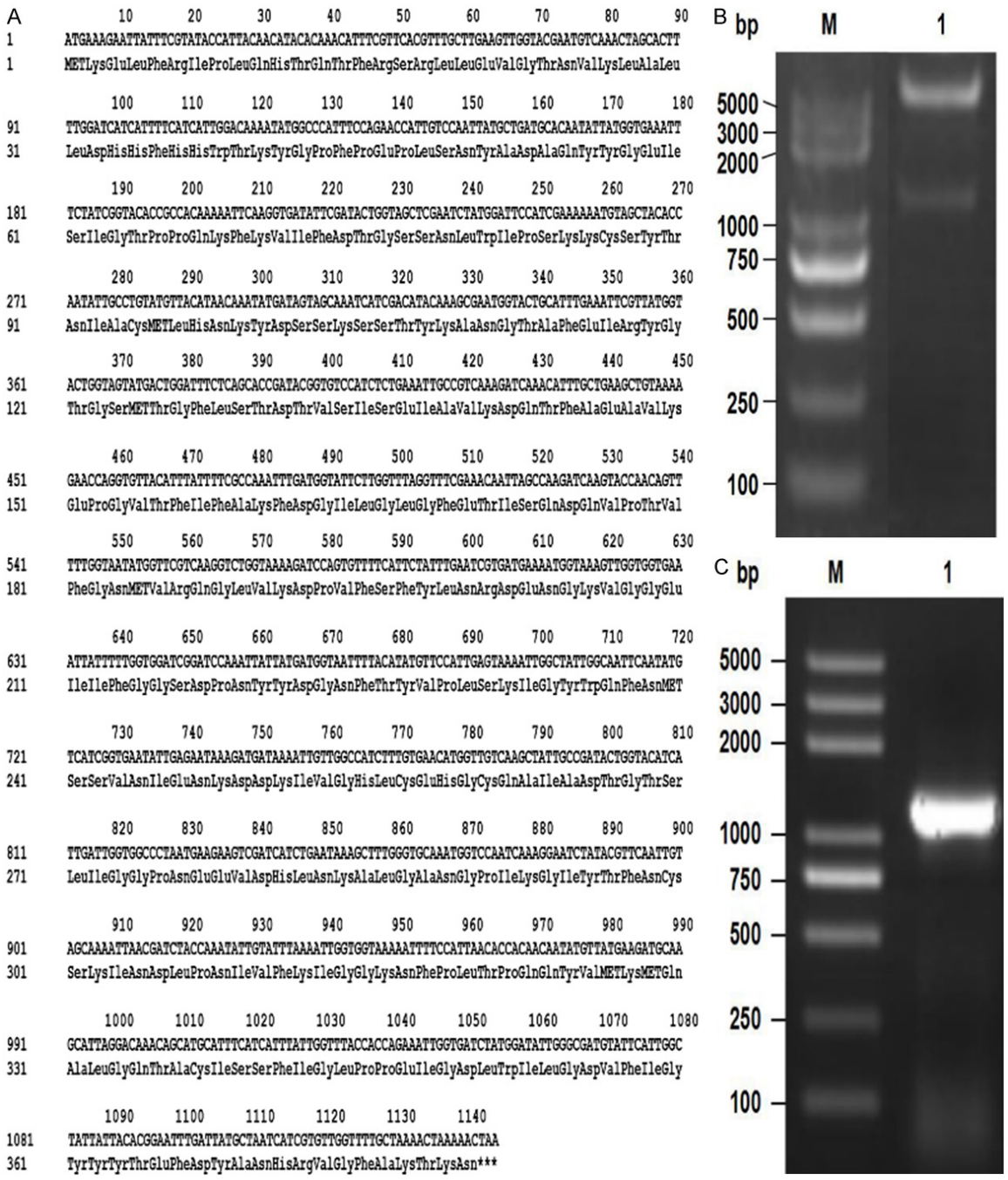


Figure 1. Cloning of cathepsin D. A. Amino acid sequences of cathepsin D. B. PCR of cathepsin D cDNA. M: DNA marker; 1: Cathepsin D gene. C. Restriction enzymatic digestion of recombinant plasmid. (M: DNA marker; 1: cathepsin D DNA).

trophoresis (SDS-PAGE). Constituents with the target protein were treated with cracking, lysis, and ultrasonication. Clear liquid on samples was then purified by the balanced Ni+ NTA column at a speed of 2 mL/min. Protein was then collected and purified by molecular sieve chromatography after adequate washing and slow elution.

Analysis of allergenicity using enzyme-linked immune sorbent assay (ELISA)

The allergenicity of cathepsin D protein was evaluated by ELISA. For antigen coating, a carbonate buffered solution (CBS. 100 μL/well, pH 9.5) was added into a microtiter plate overnight at 4°C. Next, 3% bovine serum albumin (BSA,

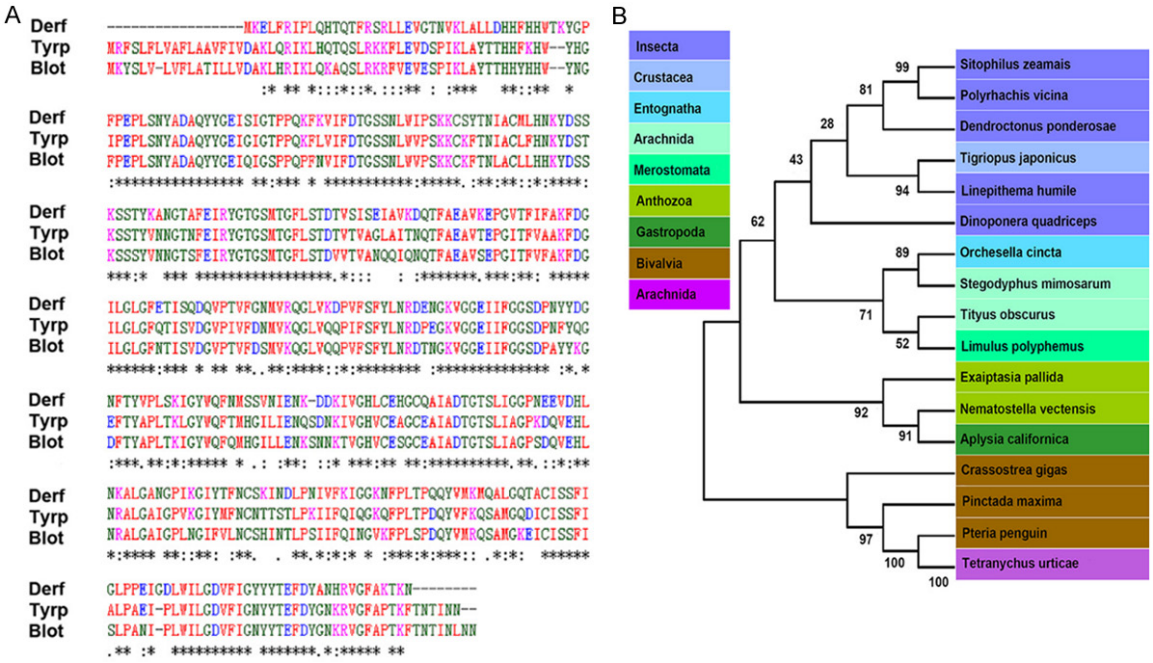


Figure 2. Sequence alignment and phylogenetic tree analysis of *D. farina* cathepsin D. A. Sequence alignment of cathepsin D from different sources. Derf (*Dermatophagoides farinae*), Tyrp (Tyrophagus putrescentiae), Blot (Blomia tropicalis). B. Phylogenetic tree of *D. farina* cathepsin D with 17 other species.

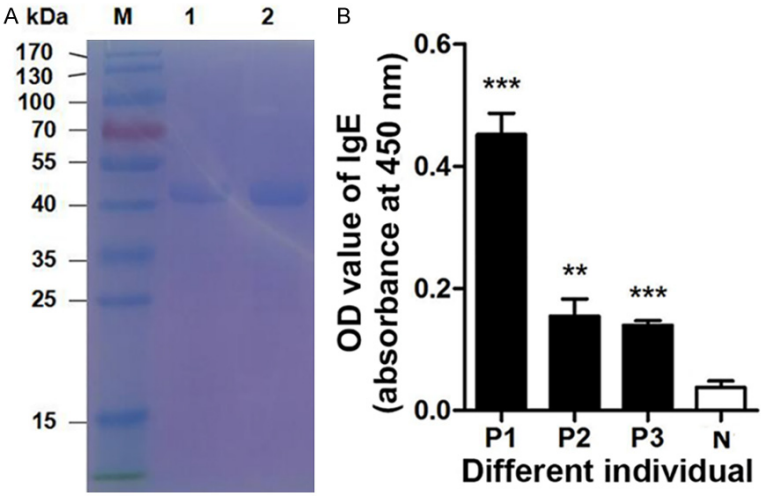


Figure 3. Preparation of recombinant cathepsin D and immunological characterization. A. SDS-PAGE analysis of the protein expressed from pET-28a (+)-cathepsin D recombinant plasmids in *E. coli*. M: Protein marker; 1: Target protein. B. Specific IgE reactivity to cathepsin D was evaluated by ELISA. P1-P3, the serum from cathepsin D SPT positive patients; N, the serum from healthy subjects. IgE levels of each patient group (P1-P3) compared to the normal group (N). t-test was used to assess statistical significance (**, P<0.05. ***, P<0.01).

200 μL/well) was used for blocking for 2 hours. Diluted asthmatic patient sera or BSA (negative control) were supplemented and incubated for 1 hour at room temperature. The plate was then incubated with peroxidase-labeled goat anti-human IgE (1:2000) for 1 hour at 37°C, followed by the addition of PBST for washing (3 times). For enzyme reactions, tetramethylbenzidine (TMB, 100 uL/well) was added and stopped by 2M H_2SO_4 (50 μL/well). Fluorescence intensity of each well was analyzed by an ELx808 absorbance microplate reader (BioTek, Shanghai, China) at 450 nm.

Skin prick test (SPT) of cathepsin D

Recombinant cathepsin D was purified and dissolved in a phosphate buffer (50 mM PB, 100 mM NaCl, pH 7.4). Sodium chloride (0.9%) and histamine (5 mg/mL) were used as negative and positive controls, respectively. Skin response was observed at 20 minutes. It was defined as positive when the prick spot became a wheal and fleck surround-

Table 1. Demographic data of mite allergic patients and skin test results of cathepsin D

Subject	Gender/Age	Diagnosis	Net wheal size (mm), Level			
			DME	Histamine	PS	r-Cathepsin D
1	Female/9	AR, FA	7, +++	5.5	0	0
2	Female/72	BA, DA	2, +	5	0	0
3	Male/17	BA, AR	14, +++	7.5	0	0
4	Female/28		2, +	4.5	0	0
5	Male/62	AR, AD, DA	2.25, +	4.75	0	0
6	Female/84	BA, FA	2.5, ++	3.5	0	0
7	Male/11	BA, AR	9, +++	5.5	0	2, +
8	Male/47	BA, AR	3.5, ++	5.5	0	0
9	Male/45	AR	2, +	7.5	0	0
10	Female/52	AR, FA	1.5, +	5.5	0	0
11	Male/20	BA, FA	1.5, +	5	0	0
12	Female/20	BA	2, ++	4	0	0
13	Female/13	AR	1.25, ++++	6	0	2, +
14	Female/66	BA, AR, DA	9, +++	5.5	0	0
15	Female/64	BA	2.25, +	6	0	0
16	Female/41	BA, AR	3.5, ++	5.5	0	0
17	Female/44	BA, AR	4.5, +++	4.5	0	1.5, +

BA (Bronchial asthma); AR (Allergic rhinitis); AD (Atopic dermatitis); FA (Food allergy); DA (Drug allergy); PS (physiological saline). Positive: ≥ 1; Negative: 0.

ing the wheal. Degrees of allergic reaction were defined as follows: 4+, response was stronger than the histamine group; 3+, response was almost the same as the histamine group; 2+, response was weaker than histamine group, but stronger than the negative group; and 1+, response was significantly weaker than histamine group, but slightly stronger than negative group.

This study was approved by the Ethics Committee and written informed consent was obtained from all participants for the use of these clinical materials for research purposes and skin tests. There were no children in the SPT.

Bioinformatics analysis of cathepsin D

After analyzing the open reading frame of cathepsin D, the amino acid sequence of the protein was deduced using DNAMAN 8 software. B-cell epitopes of cathepsin D were predicted by three bioinformatics tools [DNAStar, BepiPred 1.0 server, and BPAP (bioinformatics predicted antigenic peptides)] [22-24]. In addition, T-cell epitopes were predicted using four software/databases (SYFPEITHI, NetMHCII, NetMHCIIpan, and Preprod) in accordance to previously reported methods [22-24]. For SignalP-NN of cathepsin D prediction, SignalP 4.1 software n (http://www.cbs.dtu.dk/services/SignalP/) was used. SignalP is powerful software widely used to predict the presence and location of signal peptide sites in amino acid sequences from different organisms via several artificial neural networks [25-27]. Physicochemical properties of cathepsin D were predicted by ProtParam (http://web.expasy.org/protparam/), widely used software which is able to compute various physical and chemical parameters based on the input of a protein sequence [28, 29]. Previously, diverse studies have reported the use of ProtScale to compute profiles produced by any amino acid scale on a selected protein. In this present study, the hydrophilicity, average flexibility, and relative mutability of cathepsin D were predicted by ProtScale (http://web.expasy.org/protscale/) [30, 31]. PSIPRED, a simple and accurate method for prediction of secondary structures incorporating two feed-forward neural networks that perform analysis of output obtained from PSI-BLAST (Position Specific Iterated-BLAST), was used to predict the secondary structure of cathepsin D [32]. In a eukaryote, functions of the proteins are usually regulated by phosphorylation and dephosphorylation [33]. Given the importance of protein phosphorylation in regulating cellular signaling, this study computed the phosphorylation site of cathepsin D by NetPho 2.0. Subcellular localization of cathepsin D was predicted by CELLO2.5 [34]. Functional sites or motifs of cathepsin D were predicted by InterPro5.0 (http://www.ebi.ac.uk/interpro/) and ScanProsite (http://prosite.expasy.org/scanprosite/), in accordance with recent studies [35-37].

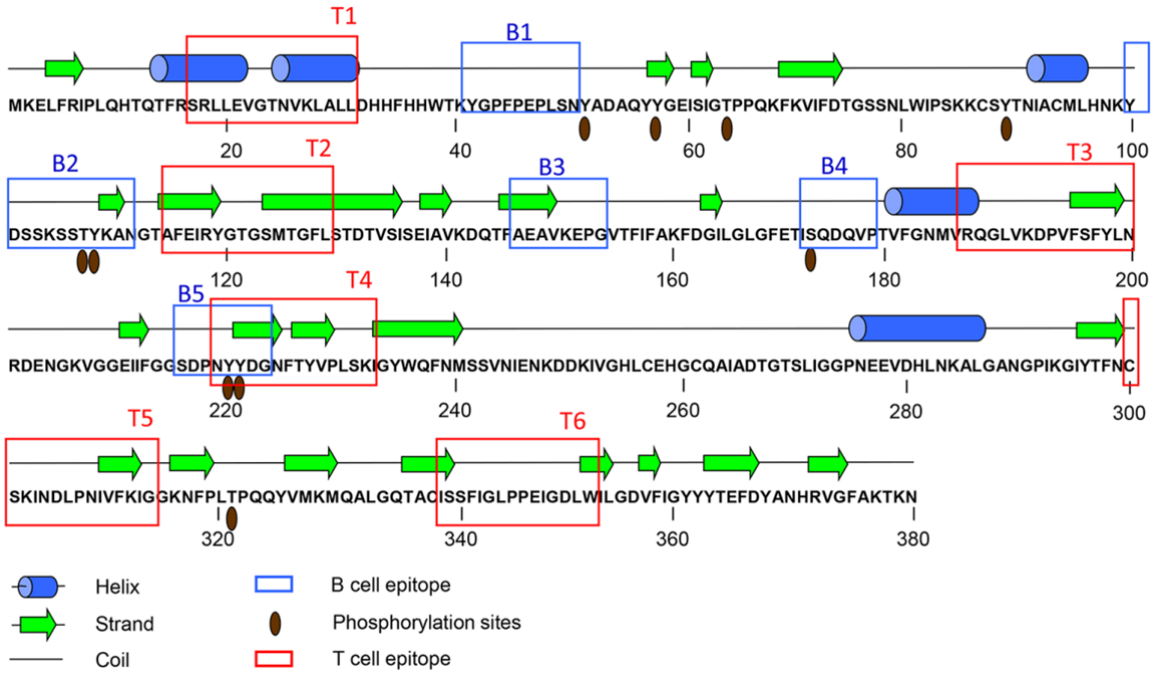


Figure 4. Secondary structure and epitope prediction. Five B-cell epitope peptides and six T-cell epitope peptides as predicted by different tools. Phosphorylation sites are also shown.

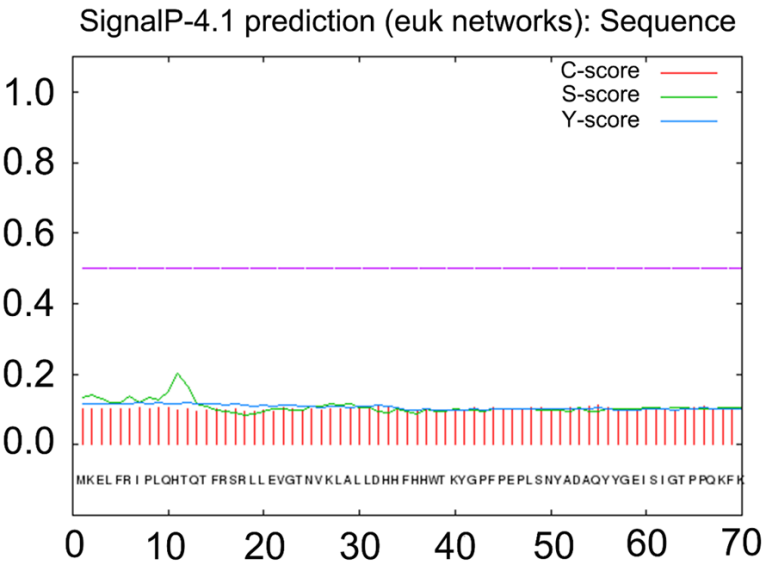


Figure 5. SignalP-NN prediction (euk networks) by SignalP 4.1 software. 'C-score' is predicted cleavage site value, 'S-score' is predicted signal peptide value, and 'Y-score' is a combination of C- and S-scores. The output figure indicates that there were no obvious peak of Y-scores. It was deduced that there was no possibility of cleavage site amino acids in cathepsin D protein.

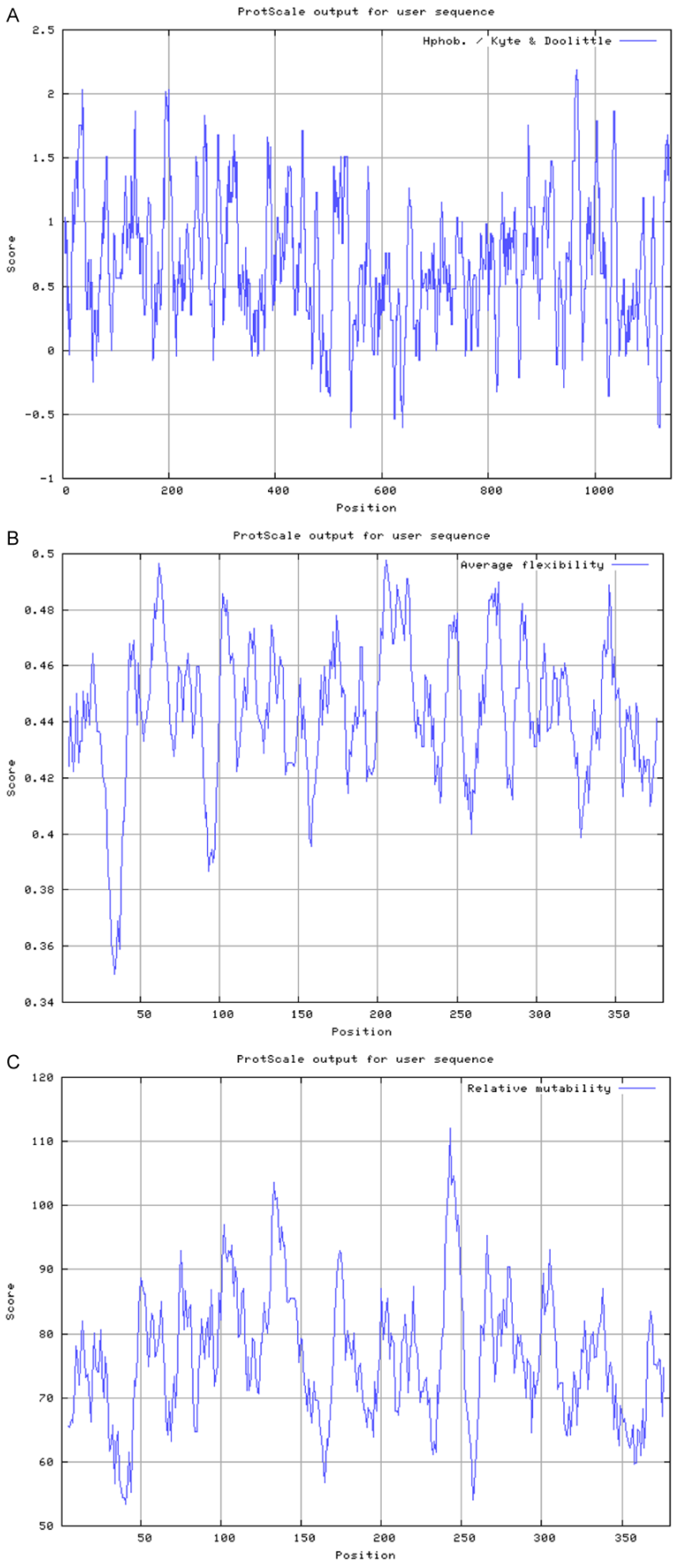
Statistical analysis

All data are expressed as mean ± SD and were processed with Graphpad software. One-way ANOVA was used for examining mean differences between patient groups (P1, P2, or P3) and the normal group (N). **, P<0.05. ***, P<0.01.

Results

Cloning and sequence alignment of the dust mite cathepsin D gene

D. farina cathepsin D genes consist of 1,143 bases, encoding 380 amino acids (**Figure 1A**). Results of agarose gel electrophoresis (**Figure 1B**) and double enzymatic digestion (**Figure 1C**) showed that the electrophoretic band of the PCR product was around 1,100, the same as the expected gene sequence of cathepsin D. This shows that the target DNA (cathepsin D) has been successfully cloned (1143 bp). Homologous comparisons (**Figure 2A**) and molecular phylogenetic analysis (**Figure 2B**) were constructed by MEGA5. 1. Results revealed that *D. farina* cathepsin D has a close relationship with *Tyrophagus putres-*



centiae, *Blomia tropicalis,* and *Sarcoptes scabiei*.

Expression and purification of recombinant cathepsin D

This study cloned the cathepsin D gene sequence into pET-28a (+), expressed, and then purified recombinant cathepsin D protein. Quality of the recombinant protein was identified by SDS-PAGE analysis (**Figure 3A**). Results showed that protein was obtained with high purity after purification.

Allergenicity analysis of cathepsin D

This study was approved and supervised by the Ethics Committee (ethical approval number: 201540, the relevant document of the experiment is supplied in Supplementary Information 1). Results of SPT (**Table 1**) showed that 3 out of 17 mite allergic patients showed positive reactions to cathepsin D. This indicates that cathepsin D has the ability of binding IgE in the serum of allergic patients. To further determine the IgE binding ability of cathepsin D, ELISA was performed using the serum of allergic patients. There were significant differences between the cathepsin D group and control group (**Figure 3**). These phenomena indicate that cathepsin D could be regarded as a subtype allergen in *D. farina*.

Structural and functional prediction

Flexibility, hydrophilicity, and surface accessibility are important properties for B-cell epitope identification. Based on these features, peptides predicted by DNAstar were 39-52, 97-111, 199-204, 216-223, 245-250, 321-325 and

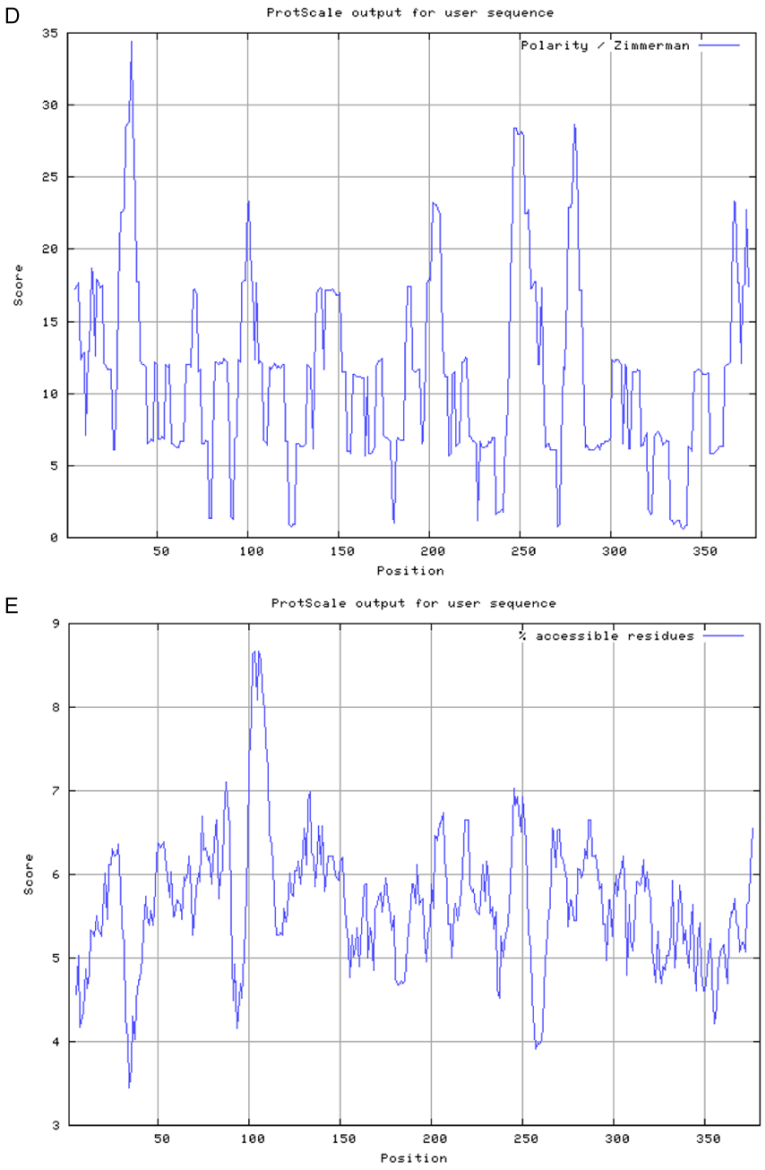


Figure 6. Bioinformatics analysis of cathepsin D. X-axis represents the protein length from N- to C-terminal. Y-axis represents the score. A. Hydrophilic predicted results of cathepsin D (ProtScale software, Kyte & Doolittle). B. Average flexibility predicted results of cathepsin D (ProtScale software). C. Relative mutability predicted results of cathepsin D (ProtScale software). D. Polarity/zimmerman of cathepsin D (ProtScale software). E. Accessible residues of cathepsin D (ProtScale software).

376-380. Peptides predicted by the Bepi-Pred 1.0 server were 40-57, 60-68, 100-112, 120-123, 143-152, 172-179, 202-208, 213-223, 245-251, 264-281, 286-293, 317-321 and 378-380. Bioinformatics-predicted antigenic peptides were 15-37, 41-50, 66-73, 79-96, 129-142, 145-169, 173-179, 186-198, 224-234, 250-263, 278-284, 305-312, 318-328, 332-347, 349-365 and 369-375. Final results of these three methods predicted 5 epitope peptides, 41-50, 100-111, 145-152, 173-179 and 216-223 (**Figure 4**). In addition, results of T-cell epitopes predictions were obtained. Six peptide sequences (aa 17-31, 114-128, 186-200, 219-233, 300-314, 338-352) were predicted as potential epitopes of cathepsin D (**Figure 4**). The secondary structure of cathepsin D predicted by PSIPRED possessed five helices and twenty-four strands (**Figure 4**). Phosphorylation sites prediction of cathepsin D with netphos 2.0 identified one serine (173), three threonine (64, 107 and 321), and six tyrosine (51, 57, 89, 108, 220 and 221) residues (**Figure 4**). SignalP-NN prediction (euk networks) was performed using SignalP4.1 software. As shown, there was no possibility of the cleavage site amino acids predicted by Signal 4.1 Server (**Figure 5**). Physicochemical properties of cathepsin D were predicted by ProtParam tools. Results suggested that protein cathepsin D, whose formula was $C_{1922}H_{2906}N_{488}O_{561}S_{13}$, consisted of 5,890 atoms, with a molecular weight of 42241.93Da. The instability index (II) of cathepsin D was computed to be 34.24. This indicated that this protein was stable. In addition, the theoretical PI, aliphatic index, grand average of hydropathicity (GRAVY) were predicted as 6.21, 77.21, and -0.222 respectively. To predict the hydrophilicity, average flexibility, relative mutability, and polarity and molar fraction of cathepsin D, ProtScale was used. As shown by output given by ProScale software scale Hphob, the protein was predicted to be highly hydrophobic (**Figure 6A**). Average flexibility results showed that the maximum score was comput-

Table 2. The prediction of protein subcellular localization (CELLO2.5) Cathepsin D

Support vector machine	Localization	Reliability
Amino Acid Comp	Extracellular	0.482
N-peptide Comp	OuterMembrane	0.680
Partitioned seq. Comp	Extracellular	0.573
Physico-chemical Comp	Extracellular	0.866
Neighboring seq. Comp	Extracellular	0.404
CELLO Prediction:		
	Extracellular	2.528
	OuterMembrane	1.131
	Periplasmic	0.920
	Cytoplasmic	0.326
	InnerMembrane	0.095

ed to be 0.498 at amino acid position 205. The minimum score was computed to be 0.350 at amino acid position 34 (**Figure 6B**). Moreover, results of relative mutability prediction showed the maximum score of 112.000 located at amino acid position 243 and minimum score of 53.333 located at amino acid position 41 (**Figure 6C**). Predicted results of polarity showed that the maximum score was 34.412 and the minimum score was 0.618, locating at amino acid positions 36 and 340, respectively (**Figure 6D**). Molar fraction (%) was computed with a maximum score of 8.667 at amino acid position 103 and minimum score of 3.456 at amino acid position 34 (**Figure 6E**). Furthermore, protein subcellular localization of cathepsin D was predicted by CELLO2.5, according to a previous study [38] (**Table 2**), with functional sites or motifs allergen of cathepsin D being computed (**Table 3**). Results of T-cell epitopes predicted were also obtained (**Table 4**). These results indicate possible physicochemical properties and structural domain of cathepsin D.

Discussion

Application of molecular biology technology to obtain new allergens can provide comprehensive information for development of SIT. There are many allergens in *D. farina* that have not been isolated and identified [12]. With the help of genome sequencing and bioinformatics, this present study has identified new allergens in *D. farina*. This should be useful for future HDM allergy diagnosis and therapy [21].

In previous studies, cathepsin D, a protein that can degrade proteins and activate precursors of bioactive proteins in pre-lysosomal compartments, has been recognized as an aspartic endo-protease widely distributed in lysosome-related organelles [13, 39]. Abnormal expression of cathepsin D can lead to several diseases, including breast cancer and possibly Alzheimer's disease [40]. Some studies have also reported that cathepsin D is involved in many pathological processes, such as inflammatory states, neoplasms, rheumatic diseases, apoptosis, and muscle dystrophy [14]. These suggest the possibility that cathepsin D may also work as an allergen in the human body.

cDNA encoding cathepsin D from *D. farina* has been successfully cloned. This study found that cathepsin D, consisting of 380 amino acids, was encoded by a gene having 1143 bp. Sequence alignment and evolutionary tree analysis showed that the homology of *D. farina* cathepsin D gene sequence with *Tyrophagus putrescentiae* cathepsin D (GenBank: AOD75-397.1) was 71%, with *Blomia tropicalis* cathepsin D (GenBank: AAX33731.1) 70%, and with *Tityus obscurus* cathepsin D (GenBank: JAT91150.1) 61%.

Allergenicity of cathepsin D was evaluated by the skin prick test, with 3 (17.6%) out of 17 house dust mite allergic patients having positive reactions to this purified protein. To further confirm the IgE-binding ability of cathepsin D, this study performed ELISA using sera from two healthy individuals and from three cathepsin D SPT positive allergic patients. Results demonstrated that cathepsin D can bind to serum IgE from cathepsin D SPT positive allergic patients. These results indicate that cathepsin D is a new allergen in *D. farina*.

Bioinformatics analysis plays an important role in predicting the structure and properties of target proteins. Sequence, structural domain, physicochemical properties, functional characteristics, and allergenticity have been predicted, enabling researchers to recognize the relationship between properties and structures. In this study, cathepsin D showed a relative stable feature with high hydrophobicity. Also, B-cell and T-cell epitopes were predicted by different tools. To improve the accuracy of com-

Table 3. Functional sites or motifs allergen

Prediction tool	Functional sites or motifs	Amino acid position
InterPro 5.0	Aspartic peptidase A1 family	3-378
	Cathepsin D	3-378
	Aspartic peptidase domain	38-138, 139-380
	Peptidase family A1 domain	56-376
	Aspartic peptidase, active site	71-82
ScanProsite	Peptidase family A1 domain profile (PEPTIDASE_A1)	56-376
	Eukaryotic and viral aspartyl proteases active site (ASP_PROTEASE)	71-82

Table 4. Epitopes prediction

Epitope	Sequence	SYFPEITHI	NetMHCII	NetMHCIIpan	Preprod
(17-31) aa	SRLLEVGTNVKLALL	+ +	+ +	+ +	- -
(114-128) aa	AFEIRYGTGSMTGFL	- -	+ +	+ +	+ +
(186-200) aa	RQGLVKDPVFSFYLN	+ +	+ +	+ +	+ +
(219-233) aa	NYYDGNFTYVPLSKI	+ -	+ -	+ +	+ +
(300-314) aa	CSKINDLPNIVFKIG	+ +	+ +	- -	+ +
(338-352) aa	ISSFIGLPPEIGDLW	+ +	+ +	+ +	+ +

putational predictions, multiple results were integrated. In addition, serial indexes including SignalP-NN, physicochemical properties, hydrophilicity, average flexibility, and relative mutability were computed by different software. Results parameters can allow for better understanding of the properties and structures of cathepsin D in *D. farina*. In conclusion, protein cathepsin D shows obvious allergenticity and physicochemical properties. The findings reported in this study should be of great help to future immunological studies.

Acknowledgements

This study was supported by grants from the Natural Science Foundation of China (No. 914-42118, 31400786, 91542104), Opening fund of State Key Laboratory of Respiratory Disease (No. SKLRD2016ZJ001), Science and Technology Project of Guangdong Province (No. 2013-B051000088, 2013B031800023, 2014B09-0901041, 2014A020212466, 2016A020216-029), Shenzhen Scientific Techology Basic Research Projects (No. JCYJ201505250929410-36, JCYJ20160328144536436, JCYJ201604-22091132854), and PhD Start-up Fund of Natural Science Foundation of Guangdong Province, China (No. 2014A030310192).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Xiaoyu Liu and Zhigang Liu, Department of Allergy, Affiliated Luohu Hospital of Shenzhen University, Shenzhen Luohu Hospital Group, Room 722 of Medical School, Shenzhen University, 3688 Nanhai Blvd, Shenzhen 518060, Guangdong, China. Tel: +86755266819-07; Fax: +8675526681906; E-mail: szulxy123@foxmail.com (XYL); lzg195910@126.com (ZGL)

References

[1] Li J, Sun B, Huang Y, Lin X, Zhao D, Tan G, Wu J, Zhao H, Cao L and Zhong N. A multicentre study assessing the prevalence of sensitizations in patients with asthma and/or rhinitis in China. Allergy 2009; 64: 1083-1092.

[2] Milián E and Díaz AM. Allergy to house dust mites and asthma. P R Health Sci J 2004; 23: 47-57.

[3] Valenta R, Campana R, Marth K and Van Hage M. Allergen-specific immunotherapy: from therapeutic vaccines to prophylactic approaches. J Intern Med 2012; 272: 144.

[4] Cui Y, Zhou Y, Wang Y, Ma G and Yang L. The group 10 allergen of Dermatophagoides farinae (Acari: Pyroglyphidae): cDNA cloning, sequence analysis, and expression in Escherichia coli BL21. J Med Entomol 2013; 50: 205-208.

[5] Shen HD, Ming FT, Huang CH, Hong C, Tai HY, Chen YS, Sheu SY and Thomas WR. Homology modeling and monoclonal antibody binding of the Der f 7 dust mite allergen. Immunol Cell Biol 2011; 89: 225-230.

[6] Kawamoto S, Mizuguchi Y, Morimoto K, Aki T, Shigeta S, Yasueda H, Wada T, Suzuki O, Jyo T and Ono K. Cloning and expression of Der f 6, a serine protease allergen from the house dust mite, Dermatophagoides farinae. Biochim Biophys Acta 1999; 1454: 201-207.

[7] King C, Simpson RJ, Moritz RL, Reed GE, Thompson PJ and Stewart GA. The isolation and characterization of a novel collagenolytic serine protease allergen (Der p 9) from the dust mite Dermatophagoides pteronyssinus. J Allergy Clin Immunol 1996; 98: 739-747.
[8] Bennett BJ and Thomas WR. Cloning and sequencing of the group 6 allergen of Dermatophagoides pteronyssinus. Clin Exp Allergy 1996; 26: 1150.
[9] Nishiyama C, Yasuhara T, Yuuki T and Okumura Y. Cloning and expression in Escherichia coli of cDNA encoding house dust mite allergen Der f 3, serine protease from Dermatophagoides farinae. FEBS Lett 1995; 377: 62-66.
[10] Smith WA, Chua KY, Kuo MC, Rogers BL and Thomas WR. Cloning and sequencing of the Dermatophagoides pteronyssinus group III allergen, Der p III. Clin Exp Allergy 1994; 24: 220.
[11] Yasueda H, Mita H, Akiyama K, Shida T, Ando T, Sugiyama S and Yamakawa H. Allergens from Dermatophagoides mites with chymotryptic activity. Clin Exp Allergy 1993; 23: 384-390.
[12] Thomas WR, Smith WA and Hales BJ. The allergenic specificities of the house dust mite. Chang Gung Med J 2004; 27: 563.
[13] Vidoni C, Follo C, Savino M, Melone MA and Isidoro C. The role of cathepsin D in the pathogenesis of human neurodegenerative disorders. Med Res Rev 2016; 36: 845-870.
[14] Minarowska A, Gacko M, Karwowska A and Minarowski Ł. Human cathepsin D. Folia Histochem Cytobiol 2008; 46: 23-38.
[15] Arruda LK, Vailes LD, Mann BJ, Shannon J, Fox JW, Vedvick TS, Hayden ML and Chapman MD. Molecular cloning of a major cockroach (Blattella germanica) allergen, Bla g 2. Sequence homology to the aspartic proteases. J Biol Chem 1995; 270: 19563-19568.
[16] Chen R and Snyder M. Promise of personalized omics to precision medicine. Wiley Interdiscip Rev Syst Biol Med 2013; 5: 73-82.
[17] Orli B. Precision medicine. Nature 2015; 526: 335.
[18] Ghaffarinazari H, Tavakkolafshari J, Jaafari MR, Tahaghoghihajghorbani S, Masoumi E and Jalali SA. Improving multi-epitope long peptide vaccine potency by using a strategy that enhances CD4+ T Help in BALB/c mice. PLoS One 2015; 10: e0142563.
[19] Navid N, Younes G, Gholamreza J, Mohammad Javad K and Eskandar O. A novel multi-epitope peptide vaccine against cancer: an in silico approach. J Theor Biol 2014; 349: 121-134.
[20] Shahsavandi S, Ebrahimi MM, Sadeghi K and Mahravani H. Design of a heterosubtypic epitope-based peptide vaccine fused with hemokinin-1 against influenza viruses. Virol Sin 2015; 30: 200-207.
[21] Chan TF, Ji KM, Yim AK, Liu XY, Zhou JW, Li RQ, Yang KY, Li J, Li M and Law PT. The draft genome, transcriptome, and microbiome of Dermatophagoides farinae reveal a broad spectrum of dust mite allergens. J Allergy Clin Immunol 2015; 135: 539.
[22] Chen H, Yang HW, Wei JF and Tao AL. In silico prediction of the T-cell and IgE-binding epitopes of Per a 6 and Bla g 6 allergens in cockroaches. Mol Med Rep 2014; 10: 2130.
[23] Li X, Yang HW, Chen H, Wu J, Liu Y and Wei JF. In Silico prediction of T and B cell epitopes of Der f 25 in Dermatophagoides farinae. Int J Genomics 2014; 2014: 483905.
[24] Yang H, Hao C, Min J, Hua X, He S and Wei JF. Molecular cloning, expression, IgE binding activities and in silico epitope prediction of Per a 9 allergens of the American cockroach. Int J Mol Med 2016; 38: 1795.
[25] Emanuelsson O, Brunak S, Von HG and Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2007; 2: 953.
[26] Petersen TN, Brunak S, Von HG and Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 2011; 8: 785-786.
[27] Zhang SW, Zhang TH, Zhang JN and Huang Y. Prediction of signal peptide cleavage sites with subsite-coupled and template matching fusion algorithm. Molecular Informatics 2014; 33: 230-239.
[28] Lafarga T, O'Connor P and Hayes M. Identification of novel dipeptidyl peptidase-IV and angiotensin-I-converting enzyme inhibitory peptides from meat proteins using in silico analysis. Peptides 2014; 59: 53-62.
[29] Thirugnanasambantham K, Muralidaran S and Mandal AK. Molecular cloning, computational and expression analysis of anthocyanidin reductase in tea (Camellia sinensis). Appl Biochem Biotechnol 2014; 174: 130.
[30] Lu YJ, Lu G, Shi DZ, Li LH and Zhong SF. Bioinformatic analysis for structure and function of TCTP from spirometra mansoni. Asian Pac J Trop Med 2013; 6: 709.
[31] Seddigh S. Comprehensive comparison of two protein family of P-ATPases (13A1 and 13A3) in insects. Comput Biol Chem 2017; 68: 266.
[32] Mcguffin LJ, Bryson K and Jones DT. The PSIPRED protein structure prediction server. Bioinformatics 2000; 16: 404-405.
[33] Hubbard MJ and Cohen P. On target with a new mechanism for the regulation of protein phosphorylation. Trends Biochem Sci 1993; 18: 172-177.
[34] Yu CS, Lin CJ and Hwang JK. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on

n -peptide compositions. Protein Sci 2004; 13: 1402.
[35] Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztányi Z, Elgebali S and Fraser M. InterPro in 2017-beyond protein family and domain annotations. Nucleic Acids Res 2017; 45: D190.
[36] Rodrigues BN, Steffens MB, Raittz RT, Santos-Weiss IC and Marchaukoski JN. Quantitative assessment of protein function prediction programs. Genet Mol Res 2015; 14: 17555.
[37] de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A and Hulo N. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 2006; 34: 362-365.
[38] Yubao C, Zhou Y, Weihong S, Guifang M, Yang L and Yungang W. Cloning, expression, and analysis of the group 2 allergen from Dermatophagoides farinae from China. An Acad Bras Cienc 2010; 82: 941-951.
[39] Diment S, Martin KJ and Stahl PD. Cleavage of parathyroid hormone in macrophage endosomes illustrates a novel pathway for intracellular processing of proteins. J Biol Chem 1989; 264: 13403.
[40] Crawford FC, Freeman MJ, Schinka J, Abdullah LI, Richards D, Sevush S, Duara R and Mullan MJ. The genetic association between Cathepsin D and Alzheimer's disease. Neurosci Lett 2000; 289: 61-65.

Supplementary Information 1. Ethics Committee for Scientific Research of the First Affiliated Hospital of Guangzhou Medical University

Clinical Trial Approval Document

Date of review: 14:45 pm, 29/7/2015

Location: Conference room of the Institute of Respiratory Diseases, twenty-ninth floor of the new building

Project number [Medical ethics permission] 2015-40, 2015-41, 2015-42, 2015-43, 2015-44, 2015-45, 2015-46, 2015-47, 2015-48

Attendance	Gender	Affiliation	Professional position or titles	Specialty	Position or duties in Ethics committee	Signature
Jingping Zheng	M	First Affiliated Hospital of Guangzhou Medical University	Deputy director of the Institute of Respiratory Diseases/ Professor	Internal medicine	Chairman	
Yuping Liu	M	First Affiliated Hospital of Guangzhou Medical University	Vice president of the First Affiliated Hospital of Guangzhou Medical University/Professor	Hospital management	Vice-chairman	
Ruliang Song	M	Guangdong Lawsons Law Office	Lawyer	Law	Committee member	
Qiaoyan Hong	F	Sub district offices of Yuexiu district in Guangzhou	Chief of the Department in the sub district offices of Yuexiu district in Guangzhou	Government regulation	Committee member	
Qianming Yao	M	First Affiliated Hospital of Guangzhou Medical University	Director of the scientific research division/Associate professor	Surgery	Committee member	None
Jinkun Huang	M	First Affiliated Hospital of Guangzhou Medical University	Vice president of the First Affiliated Hospital of Guangzhou Medical University/Associate chief physician	Surgery	Committee member	None
Guiqing Liu	F	First Affiliated Hospital of Guangzhou Medical University	Associate professor	Nursing	Committee member	
Bin Ouyang	F	First Affiliated Hospital of Guangzhou Medical University	Vice director of medical department/Chief physician	Surgery	Committee member	
Xinhua Xia	M	First Affiliated Hospital of Guangzhou Medical University	Vice director of the scientific research division/Associate professor	Traditional Chinese Medicine	Committee member	
Yao Wei	F	First Affiliated Hospital of Guangzhou Medical University	Vice director of the Pharmaceutical division/Chief pharmacist	Pharmacy	Committee member	
Zhijian Tan	M	First Affiliated Hospital of Guangzhou Medical University	Vice director of the equipment department/Senior Engineer	Medical apparatus and instruments	Committee member	
			Memberships: 11		Attendance: 9	

Remarks: none

Ethics Committee for scientific research of the First Affiliated Hospital of Guangzhou Medical University
29/7/2015

Address: No. 151 Yanjiang West road, Yuexiu district, Guangzhou 510120, Guangdong province, China. TEL: +86020-83062938.