



**BRAZILIAN JOURNAL**  
OF MEDICAL AND BIOLOGICAL RESEARCH

www.bjournal.com.br

ISSN 0100-879X  
Volume 45 (8) 681-791 August 2012

BIOMEDICAL SCIENCES

**Braz J Med Biol Res, August 2012, Volume 45(8) 746-752**

doi: 10.1590/S0100-879X2012007500077

## Cloning, bioinformatics analysis, and expression of the dust mite allergen Der f 5 of Dermatophagoides farinae

Yubao Cui, Ying Zhou, Guifang Ma, Li Yang, Yungang Wang and Weihong Shi

The Brazilian Journal of Medical and Biological Research is partially financed by



Ministério da Ciência e Tecnologia



Ministério da Educação



### Institutional Sponsors



Explore High - Performance MS Orbitrap Technology In Proteomics & Metabolomics



All the contents of this journal, except where otherwise noted, is licensed under a [Creative Commons Attribution License](https://creativecommons.org/licenses/by-nc/4.0/)

# Cloning, bioinformatics analysis, and expression of the dust mite allergen Der f 5 of *Dermatophagoides farinae*

Yubao Cui, Ying Zhou, Guifang Ma, Li Yang, Yungang Wang  
and Weihong Shi

Department of Laboratory Medicine, Yancheng Health Vocational and Technical College,  
Jiangsu Yancheng, China

## Abstract

Crude extracts of house dust mites are used clinically for diagnosis and immunotherapy of allergic diseases, including bronchial asthma, perennial rhinitis, and atopic dermatitis. However, crude extracts are complexes with non-allergenic antigens and lack effective concentrations of important allergens, resulting in several side effects. *Dermatophagoides farinae* (Hughes; Acari: Pyroglyphidae) is one of the predominant sources of dust mite allergens, which has more than 30 groups of allergen. The cDNA coding for the group 5 allergen of *D. farinae* from China was cloned, sequenced and expressed. According to alignment using the VECTOR NTI 9.0 software, there were eight mismatched nucleotides in five cDNA clones resulting in seven incompatible amino acid residues, suggesting that the Der f 5 allergen might have sequence polymorphism. Bioinformatics analysis revealed that the matured Der f 5 allergen has a molecular mass of 13604.03 Da, a theoretical pI of 5.43 and is probably hydrophobic and cytoplasmic. Similarities in amino acid sequences between Der f 5 and allergens of other domestic mite species, viz. Der p 5, Blo t 5, Sui m 5, and Lep d 5, were 79, 48, 53, and 37%, respectively. Phylogenetic analysis indicated that Der f 5 and Der p 5 clustered together. Blo t 5 and Ale o 5 also clustered together, although *Blomia tropicalis* and *Aleuroglyphus ovatus* belong to different mite families, viz. Echimyopodidae and Acaridae, respectively.

Key words: *Dermatophagoides farinae* (Hughes); Der f 5; Cloning; Expression; Bioinformatics

## Introduction

The dust mites, which are mostly represented by *Dermatophagoides* spp (Acari: Pyroglyphidae), are the major sources of indoor allergens associated with the occurrence of asthma, rhinitis, dermatitis, and other allergic diseases throughout the world (1-3). The investigation of crude dust mite extracts has demonstrated that more than 30 groups of components can induce IgE antibodies, and groups 1 and 2 were reported for the major allergens that reacted with 80% of sera from mite-sensitive patients (4,5). Of the 23 groups of dust mite allergens listed in the International Union of Immunological Societies (IUIS) nomenclature dataset (<http://www.allergen.org/>), 21 have been identified from *Dermatophagoides* spp.

Currently, crude dust mite extracts are used for the diagnosis and therapy of mite-allergic patients. The dozens of complex components in these extracts - such as allergens, non-allergic, or toxic protein, and some enzymes - are

postulated to be responsible for a variety of local and systemic pathological effects (6-8). To develop new diagnostic reagents and therapeutic strategies, a few groups of dust mite allergens have been cloned and expressed as genetic products, which were demonstrated to be more effective than their natural products (9,10).

Polymorphisms have been described in several dust mite allergens in various regions. This could influence their biological effectiveness in diagnostics and immunotherapy (11). Thus, it is necessary to characterize dust mite allergens and establish a high expression system to produce these allergens on a large scale. Group 5 allergens were listed in 30% of mid-potency dust mite allergens (12). A large amount of these allergens is necessary for therapeutic purposes as well as for diagnosis. Here, we report the cloning, sequencing and expression in *Escherichia coli* of Der f 5 cDNA from *D. farinae* isolated in China.

Correspondence: Yubao Cui, Department of Laboratory Medicine, Yancheng Health Vocational and Technical College, Jiefangnan Road, 263, Yancheng 224006, Jiangsu Province, China. E-mail: ybcui1975@hotmail.com

Received December 5, 2011. Accepted May 3, 2012. Available online May 18, 2012. Published August 3, 2012.

## Material and Methods

### Preparation of Der f 5 cDNA and polymerase chain reaction (PCR)

As described in previous reports (13-16), *D. farinae* isolated from the floors of rice and flour shops in Shanghai City, China, were cultured and isolated. About 600 mites were chosen for total RNA isolation by using an RNA isolator (Code D312; TaKaRa Biotech, China) and stored at -80°C. Based on the published sequence of Der f 5 (GenBank AY 283283), nucleic acid primers were designed and synthesized; the forward primer was 5'-GGATCCATGAAATTCATCATTGCTAT-3' and the reverse primer was 5'-CTCGAGTCAAACCTTCAATCTTTTAAACACG-3', with *Bam*HI and *Xho*I sites, respectively, at their 5'-end (underlined). First, reverse transcription (RT) was performed using the total RNA isolated from mites with the High Fidelity PrimeScript™ RT-PCR Kit (DR027; TaKaRa Biotech) in the PCR Thermal Cycler Dice (TP600; TaKaRa Biotech). Second, the RT product was used as the template for PCR in the same thermal cycler with PrimeSTAR® HS DNA Polymerase (DR010A; TaKaRa Biotech). Finally, 5 µL of the PCR product was analyzed by 1.0% agarose gel electrophoresis and visualized with ImageMaster® VDS.

### Cloning and DNA sequencing

Amplified PCR products were purified by gel electrophoresis on 1% low melting agarose gel. After the DNA was recovered with the Agarose Gel DNA Purification Kit 2.0 (DV805; TaKaRa Biotech), the purified DNA was added to a poly 'A' tail with a DNA A-Tailing Kit (D404; TaKaRa Biotech) and linked to the pMD19-T simple vector (D104; TaKaRa Biotech). The recombinant plasmid pMD19-T-Der f 5 was then transformed into *E. coli* JM109 (D9052; TaKaRa Biotech). The positive clones were screened by blue-white selection on Luria-Bertani (LB) plates containing 100 µg/mL ampicillin, and confirmed by restriction enzyme analysis with *Bam*HI and *Xho*I. DNA sequences were determined using an ABI PRISM™ 377XL DNA Sequencer (TaKaRa Biotech).

### Construction of expression plasmids

After sequencing, the recombinant plasmid was subjected to restriction enzyme digestion to release the Der f 5 fragment. At the same time, pET28a(+) (Novagen, Kit Lot No. N72770, USA) was digested with *Bam*HI and *Xho*I. The recovered Der f 5 cDNA fragment was then sub-cloned into the recovered fragment from the digested expression vector pET28a(+) to create pET28a(+)-Der f 5 using the DNA Ligation Kit (D6023; TaKaRa Biotech). The *E. coli* competent cells JM109 (D9052; TaKaRa Biotech) were transformed with pET28a(+)-Der f 5 plasmids, and positive clones were selected by blue-white screening and characterized by restriction enzyme analysis with *Bam*HI and *Xho*I.

### Expression of recombinant Der f 5 (rDer f 5) in *E. coli* BL21 (DE3)

A 0.5-µL amount of the pET28a(+)-Der f 5 plasmid was prepared using the MiniBEST Plasmid Purification Kit 2.0 (DV801A; TaKaRa Biotech) and used to transform 100 µL *E. coli* BL21 (DE3, Stratagene, USA). The *E. coli* BL21 carrying pET28a(+)-Der f 5 was grown on LB plates containing 50 µg/mL kanamycin at 37°C overnight. The single colony was inoculated into 5 mL LB containing kanamycin and then cultured at 37°C; 50 µL 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to induce the *tac* promoter. The *E. coli* cells were harvested by centrifugation and PBS (200 µL/tube) was added. After resuspension, cells were disrupted ultrasonically. Fifty microliters of lysate was taken as the whole cell lysate for later gel electrophoresis. The remaining amount was centrifuged for the separation of supernatant and pellet, regarded as soluble and insoluble protein samples, respectively. Then, 10 µL of the whole cell lysate and soluble and insoluble protein samples were subjected to SDS-PAGE with 12.5% polyacrylamide gel and CBB-R250 staining.

### Western blotting

For Western blotting analysis, proteins were transferred to a polyvinylidene fluoride membrane (Tiangen Biotech, China) with blocking buffer (1.5% BSA, 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% Tween 20) at 48 mA for 80 min. The membrane was incubated with primary antibody to Penta-His Antibody (Qiagen, Germany) 1:1000, at 4°C overnight, followed by incubation with secondary antibody of HRP-Rabbit Anti-Mouse IgG(H+L) (Zymed, USA) 1:1000 for 1 h at room temperature, and proteins were then visualized using 1 mL TrueBlue peroxidase substrate (Kirkegaard and Perry, USA) for 1 min.

### Bioinformatics of the cloned Der f 5

The open reading fragment (ORF) was determined using the ORF finder in the National Center for Biotechnology Information (NCBI) website. The amino acid sequence of Der f 5 was determined with Translate Tools on the ExPaSy webserver, its physicochemical properties were determined by the ProtParam tool, its signal peptide sequence by the SignalP 3.0 software, its hydrophilicity by ProtScale tools, the protein subcellular localization by CELLO v.2.5, and the secondary structure by the self-optimized prediction method with alignment (SOPMA) on the NPS@server.

The homology between the deduced amino acid sequence for the complete pre-pro form Der f 5 from China and other proteins was determined by comparing their sequences in all non-redundant GenBank CDS translations + PDB + SwissProt + PIR + PRF excluding environmental samples using BLASTp at the NCBI website. Based on the BLASTp search, the similar amino sequences of other mite species were chosen, i.e., *Blomia tropicalis* (Blo t 5, GenBank ABH06352), *D. pteronyssinus* (DER p 5, Gen-



Bank P14004), *Suidasia medanensis* (Sui m 5, GenBank AAX34051), *Lepidoglyphus destructor* (Lep d 5, GenBank Q9U5P2), and *Aleuroglyphus ovatus* (Ale o 5, GenBank AAX34060). Sequences were matched with those of the group 5 allergens of these other mites using ClustalW 2.0 at the European Bioinformatics Institute (EBI) website and computed for similarity with the VECTOR NTI 9.0 software (IBI, USA). The polygenetic tree was constructed for the group 5 allergens of Der f 5 and the other mite species cited above using the Mega 4.0 software.

**Results**

**Construction of the recombinant plasmids pMD19-T-Der f 5**

After extraction of the total RNA of *D. farinae*, the Der f 5 gene was synthesized by RT-PCR, and a strap of about 399 bp was present as determined by agarose gel electrophoresis. The PCR product was recovered and linked into the pMD19-T simple vector. Eight recombinant plasmids of pMD19-T-Der f 5 were subjected to enzyme digestion analysis, and the expected result was observed by 1.0% agarose gel electrophoresis (Figure 1).

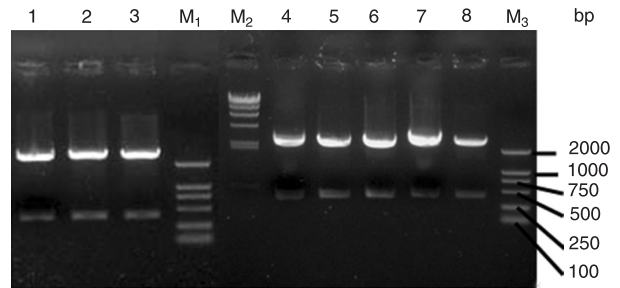
**Nucleic acid sequence and analysis**

Five of the eight plasmids determined by restriction enzyme analysis were chosen and sequenced. After removal of the vector sequence and the added restriction sites, the length of the Der f 5 nucleic acid sequence was 399 bp from the start codon ATG to the stop codon TGA. Figure 2 shows the results of alignment of our sequencing for the five PCR clones, the reference (GenBank AY283283), and another published Der f 5 sequence (GenBank AB195581) by using the VECTOR NTI 9.0 software. Eight mutation positions were observed at 35, 120, 141, 159-161, 207, 215-217, 345, and 368 bp. When these nucleotide sequences were translated into amino acid sequences, there were seven incompatible amino acid positions at 12, 40, 53, 73, 74, 115, and 123 (Figure 3). However, the nucleotide sequence for

clone 2 was 97.8% homologous to the reference.

**Construction of plasmids pET28a(+)-Der f 5 and its expression in *E. coli* BL21**

Based on sequence alignment, clone 2 was chosen for sub-cloning and expressed. After digestion of the recombinant plasmid pMD19-T-Der f 5 by both *Bam*HI and *Xho*I, the recovered cDNA encoding Der f 5 was sub-cloned



**Figure 1.** Restriction enzyme analysis of the recombinant plasmids pMD19-T-Der f 5 by *Bam*HI and *Xho*I. Lanes *M*<sub>1</sub> and *M*<sub>3</sub> = DL-2000 DNA marker; lane *M*<sub>2</sub> =  $\lambda$ -*Hind* III DNA marker; lanes 1-8 = recombinant plasmids digested with *Bam*HI and *Xho*I.

		1		100
AY283283	(1)	ATGAAATTCATCA TTGC TATTGCTGT TTGC ACTT CCGCC GTTG TATGCGTT TCGGG TGAACCGAA AAAAAACAT GATT ATCAA AATGAATT TGA TTTCT TC		
clone_1	(1)	ATGAAATTCATCA TTGC TATTGCTGT TTGC ACTT CCGCC GTTG TATGCGTT TCGGG TGAACCGAA AAAAAACAT GATT ATCAA AATGAATT TGA TTTCT TC		
clone_2	(1)	ATGAAATTCATCA TTGC TATTGCTGT TTGC ACTT CCGCC GTTG TATGCGTT TCGGG TGAACCGAA AAAAAACAT GATT ATCAA AATGAATT TGA TTTCT TC		
clone_3	(1)	ATGAAATTCATCA TTGC TATTGCTGT TTGC ACTT CCGCC GTTG TATGCGTT TCGGG TGAACCGAA AAAAAACAT GATT ATCAA AATGAATT TGA TTTCT TC		
clone_4	(1)	ATGAAATTCATCA TTGC TATTGCTGT TTGC ACTT CCGCC GTTG TATGCGTT TCGGG TGAACCGAA AAAAAACAT GATT ATCAA AATGAATT TGA TTTCT TC		
clone_5	(1)	ATGAAATTCATCA TTGC TATTGCTGT TTGC ACTT CCGCC GTTG TATGCGTT TCGGG TGAACCGAA AAAAAACAT GATT ATCAA AATGAATT TGA TTTCT TC		
AB195581	(1)	ATGAAATTCATCA TTGC TATTGCTGT TTGC ACTT CCGCC GTTG TATGCGTT TCGGG TGAACCGAA AAAAAACAT GATT ATCAA AATGAATT TGA TTTCT TC		
		101		200
AY283283	(101)	TCAT GCAAC GTAT CGAT GAC A GATGAGAA AAGGAGAG AACCATTG TTAC ATCTT CA CCA AATTA ANCA CATT CGAAGAAA COCA ACGAANGAGAT		
clone_1	(101)	TCAT GCAAC GTAT CGAT GAC A GATGAGAA AAGGAGAG AACCATTG TTAC ATCTT CA CCA AATTA ANCA CATT CGAAGAAA COCA ACGAANGAGAT		
clone_2	(101)	TCAT GCAAC GTAT CGAT GAC A GATGAGAA AAGGAGAG AACCATTG TTAC ATCTT CA CCA AATTA ANCA CATT CGAAGAAA COCA ACGAANGAGAT		
clone_3	(101)	TCAT GCAAC GTAT CGAT GAC A GATGAGAA AAGGAGAG AACCATTG TTAC ATCTT CA CCA AATTA ANCA CATT CGAAGAAA COCA ACGAANGAGAT		
clone_4	(101)	TCAT GCAAC GTAT CGAT GAC A GATGAGAA AAGGAGAG AACCATTG TTAC ATCTT CA CCA AATTA ANCA CATT CGAAGAAA COCA ACGAANGAGAT		
clone_5	(101)	TCAT GCAAC GTAT CGAT GAC A GATGAGAA AAGGAGAG AACCATTG TTAC ATCTT CA CCA AATTA ANCA CATT CGAAGAAA COCA ACGAANGAGAT		
AB195581	(101)	TCAT GCAAC GTAT CGAT GAC A GATGAGAA AAGGAGAG AACCATTG TTAC ATCTT CA CCA AATTA ANCA CATT CGAAGAAA COCA ACGAANGAGAT		
		201		300
AY283283	(198)	GAAA GAACA GATC TTAG CAAGT GAAA TCGA TACT ATTA TCGA CTGA TCGA CCGTG TTGG TGGT GTTT TGA TCGT CTTAT GAAA CGTA COGA TTTGGAC		
clone_1	(201)	GAAA GAACA GATC TTAG CAAGT GAAA TCGA TACT ATTA TCGA CTGA TCGA CCGTG TTGG TGGT GTTT TGA TCGT CTTAT GAAA CGTA COGA TTTGGAC		
clone_2	(201)	GAAA GAACA GATC TTAG CAAGT GAAA TCGA TACT ATTA TCGA CTGA TCGA CCGTG TTGG TGGT GTTT TGA TCGT CTTAT GAAA CGTA COGA TTTGGAC		
clone_3	(201)	GAAA GAACA GATC TTAG CAAGT GAAA TCGA TACT ATTA TCGA CTGA TCGA CCGTG TTGG TGGT GTTT TGA TCGT CTTAT GAAA CGTA COGA TTTGGAC		
clone_4	(201)	GAAA GAACA GATC TTAG CAAGT GAAA TCGA TACT ATTA TCGA CTGA TCGA CCGTG TTGG TGGT GTTT TGA TCGT CTTAT GAAA CGTA COGA TTTGGAC		
clone_5	(201)	GAAA GAACA GATC TTAG CAAGT GAAA TCGA TACT ATTA TCGA CTGA TCGA CCGTG TTGG TGGT GTTT TGA TCGT CTTAT GAAA CGTA COGA TTTGGAC		
AB195581	(201)	GAAA GAACA GATC TTAG CAAGT GAAA TCGA TACT ATTA TCGA CTGA TCGA CCGTG TTGG TGGT GTTT TGA TCGT CTTAT GAAA CGTA COGA TTTGGAC		
		301		400
AY283283	(298)	ATAT TTGAA CGAT ATAA CSTAGAAA TCCA TTGA AATCTAATGAAAT TTGGAAGG TGA TCTTA AAAAAAGAA GAAC AAGGT GTTA AAAAA GATT GAAGT TT		
clone_1	(298)	ATAT TTGAA CGAT ATAA CSTAGAAA TCCA TTGA AATCTAATGAAAT TTGGAAGG TGA TCTTA AAAAAAGAA GAAC AAGGT GTTA AAAAA GATT GAAGT TT		
clone_2	(298)	ATAT TTGAA CGAT ATAA CSTAGAAA TCCA TTGA AATCTAATGAAAT TTGGAAGG TGA TCTTA AAAAAAGAA GAAC AAGGT GTTA AAAAA GATT GAAGT TT		
clone_3	(298)	ATAT TTGAA CGAT ATAA CSTAGAAA TCCA TTGA AATCTAATGAAAT TTGGAAGG TGA TCTTA AAAAAAGAA GAAC AAGGT GTTA AAAAA GATT GAAGT TT		
clone_4	(298)	ATAT TTGAA CGAT ATAA CSTAGAAA TCCA TTGA AATCTAATGAAAT TTGGAAGG TGA TCTTA AAAAAAGAA GAAC AAGGT GTTA AAAAA GATT GAAGT TT		
clone_5	(298)	ATAT TTGAA CGAT ATAA CSTAGAAA TCCA TTGA AATCTAATGAAAT TTGGAAGG TGA TCTTA AAAAAAGAA GAAC AAGGT GTTA AAAAA GATT GAAGT TT		
AB195581	(298)	ATAT TTGAA CGAT ATAA CSTAGAAA TCCA TTGA AATCTAATGAAAT TTGGAAGG TGA TCTTA AAAAAAGAA GAAC AAGGT GTTA AAAAA GATT GAAGT TT		
		401		
AY283283	(398)	GA		
clone_1	(398)	GA		
clone_2	(398)	GA		
clone_3	(398)	GA		
clone_4	(398)	GA		
clone_5	(398)	GA		
AB195581	(398)	GA		

**Figure 2.** Nucleotide sequence alignment among different plasmids sequenced in this paper and the Der f 5 sequences published in GenBank by using the VECTOR NTI 9.0 software (IBI, USA).

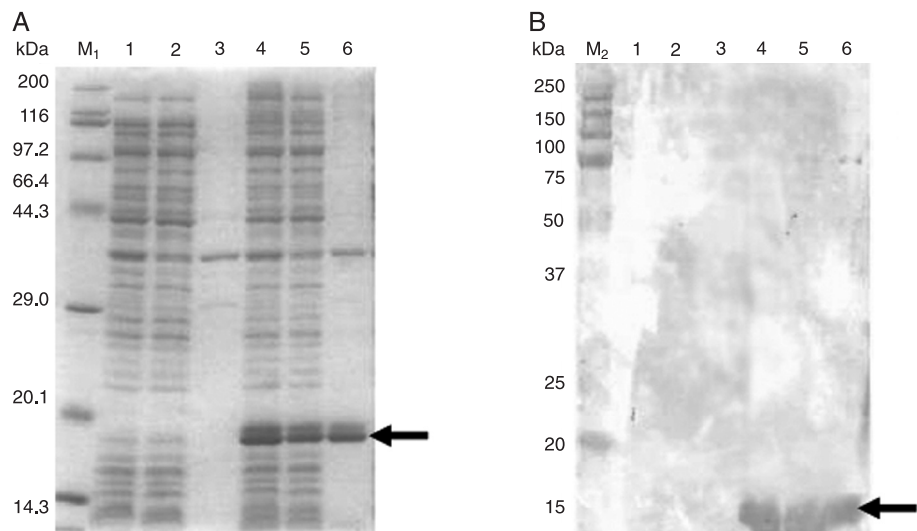
into the expression vector of pET28a(+) and identified by restriction digestion. *E. coli* BL21 was then transformed with plasmid pET28a(+)-Der f 5 and the protein expressed was induced with IPTG at 37°C for 3 h. Absorbance at 600 nm was measured for 0.562 and 1.259 before and after induction. A single specific band from SDS-PAGE and Western blotting (Figure 4A and B) was observed, corresponding to the expected molecular weight.

#### Inferred amino acid sequence and its structural and functional prediction

Five recombinant plasmids pMD19-T-Der f 5 were subjected to DNA sequencing, and the highest homology (97.8%) with the reference sequence was observed in clone 2. One ORF was identified encoding a protein of 132 amino acids, 15525.1 Da, and theoretical isoelectric point (pI) of 5.62. A signal peptide sequence of 1-19 amino acid residues was predicted by the SignalP 3.0 software. After removal of this leading sequence, the matured part of Der f 5 consisted of 113-amino acid residues (13604.03 Da, theoretical pI = 5.43). The grand average of hydropathicity (GRAVY) of -0.803 indicated that this is a hydrophobic protein, confirmed by the ProtScale software. On the NPS@server, the secondary structure of the matured Der f 5 allergen was predicted using the SOPMA software, indicating that 90.3% (102 amino acids) of the protein was alpha helix, 0.88% (1 amino acid) extended strand, 0.88% beta turn, and 7.96% (9 amino acids) random coils. Its subcellular localization was predicted by the subCELLular LOCALization predictive system (CELLO) to be cytoplasmic (Table 1). The PROSCAN sequence search tool on the NPS@server indicated that the protein contains two casein kinase II phosphorylation sites (20-23 and 58-61 residues).



**Figure 3.** Results of alignment among amino acid sequences deduced from different plasmids sequenced in this paper and the Der f 5 sequences published in GenBank by using the VECTOR NTI 9.0 software (IBI, USA).



**Figure 4.** Expression of rDer f 5 in *Escherichia coli* BL21 cells. *E. coli* BL21 cells were transformed with either pET28a(+)-Der f 5 or empty vector pET28a(+) as control. **A**, SDS-PAGE analysis of the rDer f 5 protein. Lane M<sub>1</sub> = TaKaRa protein marker (Broad); lane 1 = whole cell lysate of *E. coli* BL21 cells containing pET28a; lane 2 = supernatant of cells containing pET28a; lane 3 = pellet of cells containing pET28a; lane 4 = whole cell lysate of *E. coli* BL21 cells containing pET28a(+)-Der f 5; lane 5 = supernatant of cells containing pET28a(+)-Der f 5; lane 6 = pellet of cells containing pET28a(+)-Der f 5. **B**, Western blotting analysis of the rDer f 5 protein. Lane M<sub>2</sub> = Precision Plus Protein Standards; lane 1 = whole cell lysate of *E. coli* BL21 cells containing pET28a; lane 2 = whole cell lysate of *E. coli* BL21 cells containing pET28a(+)-Der f 5. Arrows point to the band of rDer f 5.

#### Analysis of amino acid sequence homology, alignment and molecular evolution

The similarity of Der f 5 and Der p 5, Blot 5, Sui m 5, and Lep d 5 was 79, 48, 53, and 37%, respectively (Figure 5, Table 2); Der f 5 and Der p 5 were clustered with 99% bootstrap, whereas Blot 5 and Ale o 5 were clustered with 78% (Figure 6).

#### Discussion

We cloned for the first time the full-length cDNA frag-

ment encoding the group 5 allergen from *D. farinae* isolated in China. Five sequenced cDNA clones of Der f 5 were obtained simultaneously, and there still are eight mismatched nucleotides and seven incompatible amino acid residues in the complete pre-protein. This indicates that the Der f 5 allergen may have sequence polymorphism; only a single amino acid variation (at position 61, from alanine to aspartate) was presented in the amino acid sequence deduced from five cDNA clones of Der p 5 (17). More research should be done regarding the sequence polymorphisms of the group 5 dust mite allergens, which will be important for the development of peptide-based immunotherapeutic agents for mite-allergic patients.

Among 19 patients with positive skin tests to *D. pteronyssinus* extracts, 10 were positive to rDer p 5 (18), so the group 5 allergen should be incorporated into large-scale clinical trials. Therefore, we expressed the Der f 5 allergen in *E. coli*. A single specific band between 14 and 20 kDa was observed after SDS-PAGE and Western blotting. The expression vector pET28a(+), containing a nickel-binding 6-histidine, indicates that the recombinant product has an extra 6-histidine and the molecular weight of rDer f 5 should be lower than determined by SDS-PAGE and Western blotting in the present study.

When the complete sequence of Der f 5 cDNA was analyzed, one ORF with the full length of 399 bp was found included in the cDNA fragment, which was reckoned to encode a protein with 132-amino acid residues. Further analysis showed that there was a signal peptide sequence from 1- to 19-amino acid residues. After removal of this sequence,

**Table 1.** Result of the prediction of protein subcellular localization for Der f 5 by CELLO v.2.5 (<http://cello.life.nctu.edu.tw/>).

Support vector machine	Localization	Reliability
Amino acid composition	Cytoplasm	0.357
N-peptide composition	Cytoplasm	0.993
Partitioned sequence composition	Cytoplasm	0.743
Physicochemical composition	Cytoplasm	0.937
Neighboring sequence composition	Cytoplasm	0.932
CELLO prediction	Cytoplasm	3.962
	Periplasmic	0.359
	Extracellular	0.310
	Inner membrane	0.259
	Outer membrane	0.110

**Table 2.** Analysis of similarity between Der f 5 and its homologous amino acid sequence of other mite species by the VECTOR NTI 9.0 software.

Mite protein	Similarity (%)					
	Ale o 5	Blo t 5	Der f 5	Der p 5	Sui m 5	Lep d 5
Ale o 5	-	52	37	37	50	32
Blo t 5	-	-	48	44	50	38
Der f 5	-	-	-	79	53	37
Der p 5	-	-	-	-	52	32
Sui m 5	-	-	-	-	-	39
Lep d 5	-	-	-	-	-	-

Ale o = *Aleuroglyphus ovatus*; Blo t = *Blomia tropicalis*; Der f = *Dermatophagoides farinae*; Der p = *D. pteronyssinus*; Sui m = *Suidasia medanensis*; Lep d = *Lepidoglyphus destructor*.



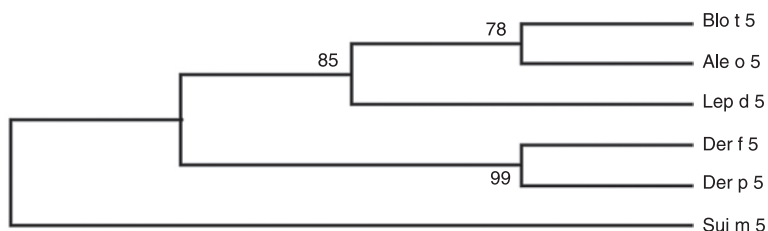
**Figure 5.** Alignment between Der f 5 and its homologous amino acid sequence of other mite species by ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/>). Asterisks = residues or nucleotides in the column are identical in all sequences in the alignment; colons = conserved substitutions; periods = semi-conserved substitutions. For abbreviations, see legend to Table 2.



the active part of Der f 5 should consist of 113-amino acid residues with a calculated molecular mass of 13604.03 Da and theoretical pI of 5.43 - identical to previous reports that identified a partial cDNA clone of Der p 5, producing a recombinant polypeptide that reacted with a 14-kDa band in mite extracts (19).

The general three-dimensional form of local segments of proteins (i.e., their secondary structure) is defined by patterns of hydrogen bonds between backbone amide and carboxyl groups. In the present study, the SOPMA software was used to predict the secondary structure of the matured part of Der f 5, and the analysis indicated that the allergen consisted of alpha helix (90.3%), extended strand (0.9%), beta turn (0.9%), and random coils (8.0%). Subsequently, the subcellular localization for Der f 5 showed that it should be cytoplasmic. However, this is a prediction, and its location must be confirmed experimentally.

Much of the published literature concerns allergens of house dust mites, which belong to the family Pyroglyphidae, especially *D. pteronyssinus*, *D. farinae*, and *Euroglyphus maynei*. Other mite species, referred to as 'storage mites', are also considered allergenic, but their study has been more limited (20). We utilized the deduced amino acid sequence of Der f 5 from China in NCBI and similar amino acid sequences of other mites were obtained, including *D. pteronyssinus* (Pyroglyphidae), *L. destructor* (Glycyphagidae) and *B. tropicalis* (Echimyopodidae), *S. medanensis* (Suidasiidae), and *A. ovatus* (Acaridae). These mite species were considered to be important from an economic and sanitary perspective and are now being recognized as important contributors to



**Figure 6.** Phylogenetic tree constructed for Der f 5 and its homologous amino acid sequence of other mite species by Mega 4.0. The values are based on bootstrap analysis (N = 6). For abbreviations, see legend to Table 2.

the allergen content of house dust (4,5,20). They can cause occupational respiratory allergies in farmers and other occupationally exposed individuals (20). Many mite allergens have shown sequence homology and biological function similar to those described in *Dermatophagoides* spp (20). Sequence similarities in the present study were 79, 48, 53, and 37% between Der f 5 and Der p 5, Blo t 5, Sui m5, and Lep d 5, respectively. In the phylogenetic tree constructed based on the group 5 allergens of the various mite species, Der f 5 and Der p 5 clustered with 99% bootstrap. Interestingly, Blo t 5 and Ale o 5 clustered with 78%, despite the fact that these two mites belong to different families. This illustrates the need for more molecular data in the study of relationships among mite species.

## Acknowledgments

Research supported by the National Sciences Foundation of China (#NSFC30060166 and #NSFC81001330) and by the Jiangsu Provincial Health Department (Grants #Z200914 and #J200907).

## References

- Milian E, Diaz AM. Allergy to house dust mites and asthma. *P R Health Sci J* 2004; 23: 47-57.
- Nadchatram M. House dust mites, our intimate associates. *Trop Biomed* 2005; 22: 23-37.
- Bunnag C, Jareoncharsri P, Tantilipikorn P, Vichyanond P, Pawankar R. Epidemiology and current status of allergic rhinitis and asthma in Thailand - ARIAAsia-Pacific Workshop report. *Asian Pac J Allergy Immunol* 2009; 27: 79-86.
- Thomas WR, Smith WA, Hales BJ, Mills KL, O'Brien RM. Characterization and immunobiology of house dust mite allergens. *Int Arch Allergy Immunol* 2002; 129: 1-18.
- Thomas WR, Smith WA, Hales BJ. The allergenic specificities of the house dust mite. *Chang Gung Med J* 2004; 27: 563-569.
- Dinakar C, Portnoy JM. Allergen immunotherapy in the prevention of asthma. *Curr Opin Allergy Clin Immunol* 2004; 4: 131-136.
- Mauro M, Russello M, Alesina R, Sillano V, Alessandrini A, Dama A, et al. Safety and pharmacoeconomics of a cluster administration of mite immunotherapy compared to the traditional one. *Eur Ann Allergy Clin Immunol* 2006; 38: 31-34.
- Zhang L, Wang C, Han D, Wang X, Zhao Y, Liu J. Comparative study of cluster and conventional immunotherapy schedules with *Dermatophagoides pteronyssinus* in the treatment of persistent allergic rhinitis. *Int Arch Allergy Immunol* 2009; 148: 161-169.
- Jeong KY, Hongb CS, Yong TS. Recombinant allergens for diagnosis and immunotherapy of allergic disorders, with emphasis on cockroach allergy. *Curr Protein Pept Sci* 2006; 7: 57-71.
- Mothes N, Valenta R, Spitzauer S. Allergy testing: the role of recombinant allergens. *Clin Chem Lab Med* 2006; 44: 125-132.
- Piboonpocanun S, Malainual N, Jirapongsananuruk O, Vichyanond P, Thomas WR. Genetic polymorphisms of major house dust mite allergens. *Clin Exp Allergy* 2006; 36: 510-516.
- Thomas WR, Heinrich TK, Smith WA, Hales BJ. Pyroglyphid

- house dust mite allergens. *Protein Pept Lett* 2007; 14: 943-953.
13. Cui Y, Zhou P, Peng J, Peng M, Zhou Y, Lin Y, et al. Cloning, sequence analysis, and expression of cDNA coding for the major house dust mite allergen, Der f 1, in *Escherichia coli*. *Braz J Med Biol Res* 2008; 41: 380-388.
  14. Cui YB, Cai HX, Li L, Zhou Y, Gao CX, Shi WH, et al. Cloning, sequence analysis and expression in *E. coli* of the group 3 allergen of *Dermatophagoides farinae*. *Chin Med J* 2009; 122: 2657-2661.
  15. Yu-bao C, Zhou Y, Weihong S, Guifang M, Yang L, Yungang W. Cloning, expression, and analysis of the group 2 allergen from *Dermatophagoides farinae* from China. *An Acad Bras Cienc* 2010; 82: 941-951.
  16. Cui YB, Cai HX, Zhou Y, Gao CX, Shi WH, Yu M, et al. Cloning, expression, and characterization of Der f 7, an allergen of *Dermatophagoides farinae* from China. *J Med Entomol* 2010; 47: 868-876.
  17. Lin KL, Hsieh KH, Thomas WR, Chiang BL, Chua KY. Characterization of Der p V allergen, cDNA analysis, and IgE-mediated reactivity to the recombinant protein. *J Allergy Clin Immunol* 1994; 94: 989-996.
  18. Simpson A, Green R, Custovic A, Woodcock A, Arruda LK, Chapman MD. Skin test reactivity to natural and recombinant *Blomia* and *Dermatophagoides* spp allergens among mite allergic patients in the UK. *Allergy* 2003; 58: 53-56.
  19. Tovey ER, Johnson MC, Roche AL, Cobon GS, Baldo BA. Cloning and sequencing of a cDNA expressing a recombinant house dust mite protein that binds human IgE and corresponds to an important low molecular weight allergen. *J Exp Med* 1989; 170: 1457-1462.
  20. Fernandez-Caldas E, Iraola V, Carnes J. Molecular and biochemical properties of storage mites (except *Blomia* species). *Protein Pept Lett* 2007; 14: 954-959.