

Allergens of wild house dust mites: Environmental Der p 1 and Der p 2 sequence polymorphisms

Wendy-Anne Smith, PhD, Belinda J. Hales, PhD, Andrew G. Jarnicki, BSc(Hons), and Wayne R. Thomas, PhD *West Perth, Australia*

Background: Sequence diversity is a common feature of mite allergens. Previous studies, using predominantly commercial mite clones, have described several polymorphic residues for Der p 1 and Der p 2.

Objective: This study aimed at determining the occurrence of sequence diversity in environmental mite isolates.

Methods: Mites were isolated from houses in Perth and Sydney, Australia. Total RNA was extracted from 1 to 30 Perth mites, and cDNA was synthesized by reverse transcriptase PCR. Der p 1 and Der p 2 cDNAs were PCR amplified and sequenced. Genomic Der p 1 DNA was amplified from whole Sydney mites directly by PCR and then sequenced.

Results: Twelve Der p 1 and 9 Der p 2 cDNA clones and 3 Der p 1 genomic DNA were analyzed and showed a high frequency of amino acid polymorphisms. Der p 2 displayed a clear pattern of divergence toward 2 alleles that differed by 4 amino acids and had characteristic silent nucleotide changes. The pattern for Der p 1 was different and unusual, with almost no silent nucleotide substitutions but frequent sporadic missense changes. Proliferative responses of peripheral blood mononuclear cells to peptides containing polymorphic residues of Der p 1 were detected in 8 of 19 subjects, with stimulation being found only for either one of the variant forms of the peptides. However, the responses to variants of whole recombinant allergens were similar, as shown for 4 variants of Der p 2.

Conclusion: Two clones for each of the allergens were identified as containing sequences that were largely representative of environmental isolates. A small-scale reverse transcriptase PCR used to produce cDNA from individual mites isolated from house dust will have wide application for studies on mite genetics and the production of recombinant mite allergens. Differences in T-cell responses to peptides representing variant epitopes were found, but responses to variants of whole recombinant allergens were similar. The GenBank and Swiss Prot database entries for Der p 1 (U11695) and Der p 2 (P49278) have been updated with the inclusion of the sequence polymorphisms described in this study. (J Allergy Clin Immunol 2001;107:985-92.)

Key words: Mites, allergens, environment, polymorphism, epitopes, T cells

The group 1 and 2 allergens of *Dermatophagoides pteronyssinus* and *Dermatophagoides farinae* house dust mites have been well characterized with cDNA synthesized from RNA isolated from a commercial source of mites (CSL Ltd, Melbourne, Australia). The commercial mites used to make these cDNA libraries had been kept in colonies isolated from the environment for approximately 30 years. It is important to consider therefore that the allergens synthesized by these mites may differ from the allergens produced by the "wild" mites that are found in the soft furnishings and carpets of homes and that initiate sensitization.

Amino acid sequence diversity of individual allergens from a variety of sources has been described for plant allergens in particular¹⁻⁴ but also for some house dust mite and mammalian allergens.⁵ Diversity arises either as a result of multigene families, as is often the case with plant allergens, in which the greatest degree of sequence divergence is observed, or from allelic polymorphisms. Work published by our group and others, using mites cultured for commercial extracts, has shown that intraspecies sequence variation is a common feature of the group 1 and group 2 allergens.⁶⁻¹⁰ The genes are present as single loci, so the polymorphisms are due to allelic variations. The polymorphisms occur frequently, with almost every Der p 1 and Der p 2 clone sequenced differing by between 1 and 4 residues, but because of limited information, no pattern of diversity could be inferred.^{6,7}

The occurrence and effects of allergen diversity need to be considered in the development of more appropriate allergen extracts and for new immunotherapeutic strategies such as the use of peptides or DNA vaccines. This study describes the sequence polymorphisms of the major group 1 and 2 allergens produced by environmental mites and shows that T cells can respond differently to peptides containing the different amino acid substitutions. Apart from one frequent substitution, Der p 1 exhibited sporadic amino acid changes in the absence of silent mutations, whereas Der p 2 had a pattern of sequence divergence, with both silent and missense changes.

METHODS

Materials

D pteronyssinus mites were collected in Perth, Western Australia, and by Dr Euan Tovey from Sydney. They were collected by vacuuming carpet or soft furnishings for 2 to 3 minutes with filter paper in a filter apparatus attached to the head of a vacuum cleaner pipe. The dust collected in the filter from the Perth houses was

From the TVW Telethon Institute for Child Health Research and the Centre for Child Health Research, University of Western Australia.

Supported by the Australian National Health and Medical Research Council. Received for publication June 9, 2000; revised January 25, 2001; accepted for publication January 25, 2001.

Reprint requests: Wayne R. Thomas, PhD, TVW Telethon Institute for Child Health Research, PO Box 855, West Perth, Western Australia 6872, Australia.

Copyright © 2001 by Mosby, Inc.

0091-6749/2001 \$35.00 + 0 1/83/114652

doi:10.1067/mai.2001.114652

TABLE I. Der p 2 environmental and commercial nucleotide polymorphisms

Nucleotide residue No.	Der p 2 polymorphisms			Clones containing polymorphisms		
	AA residue No.	Reference (Der p 2.0101) (triplet/AA)	Polymorphic (triplet/AA)	Environmental cDNA	CSL cDNA	CSL genomic DNA
117	39*	GCC/Ala	GCT/Ala	2	3	a, b, d
118/120	40	GTT/Val	TTA/Leu	2	3	a, b, d
132	44*	AAC/Asn	AAT/Asn	2	3	a, b, d
139	47	ACA/Thr	TCA/Ser	2, 3, 4, 5	2, 3	a, b, d
147	49*	ACG/Thr	ACA/Thr	2	3	a, b
147	49*	ACG/Thr	ACC/Thr	1, 3-8	2	c, d
168	56*	GCC/Ala	GCT/Ala	2, 3, 4, 5	2, 3	a, b, d
183	61*	TTA/Leu	TTG/Leu	—	3	—
225	75*	TAC/Tyr	TAT/Tyr	2, 3, 4, 5	2, 3	a, b, c, d
234	78*	TGC/Cys	TGT/Cys	2, 3, 4, 5	2, 3	a, b, d
243	81*	GTT/Val	GTC/Val	6	—	—
285	95*	CCG/Pro	CCA/Pro	2	2, 3	a, d
292	98	GCA/Ala	ACA/Thr	1	—*	—
324	108*	GTT/Val	GTC/Val	2, 3, 4, 5, 6	—	b
331	111	ATG/Met	TTG/Leu	2	—	a
340	114	GAT/Asp	AAT/Asn	2	2, 3	a
347	116	GTT/Val	GCT/Ala	5	—	—
354	118*	GCC/Ala	GCT/Ala	—	3	—
379	127	ATC/Ile	CTC/Leu	—	2	—

The nucleotide and corresponding amino acid numbering is based on the numbering for the reference sequence encoding the Der p 2.0101 mature protein¹⁸ (GenBank accession No. P49278). Environmental clones 1 and 8, clones 2 and 9, clones 3 and 6, and clones 4, 5, and 7 were each isolated from the same houses. The CSL cDNA and genomic DNA clones were described previously by Chua et al.⁷

AA, Amino acid.

*Silent substitutions.

placed in a mesh container secured at the top of a funnel; a 1.5-mL Eppendorf was placed at the base of the funnel and used to collect the mites after exposure of the dust to light and heat from a 60-W lamp for a period of 30 minutes. The mites collected from the Sydney houses were cultured, dried, and stored at 4°C.

Amplification of genomic DNA from wild mites

Five to eight mites were placed in each 0.2-mL PCR tube, with 50 µL of PCR master mix prepared as outlined by the manufacturer's specifications with *Pfu* polymerase (Promega Corp, Sydney, Australia). The thermocycler protocol used was as follows: 98°C, 10 minutes; [98°C, 30 seconds; 37°C, 30 seconds; 72°C, 1 minute] × 5 cycles; [98°C, 30 seconds; 50°C to 60°C, 30 seconds; and 72°C, 1 minute] × 35 cycles.

Reverse transcriptase PCR of total RNA isolated from wild mites

Total RNA was extracted from between 1 and 30 mites with an Ambion Totally RNA kit (Ambion Inc, Austin, Tex). The total RNA was prepared as outlined by the manufacturers with the following exceptions: (1) the mites were homogenized for two 30-second cycles with an Eppendorf minihomogenizer and pestle, using between 25 µL (1-10 mites) and 50 µL (10-50 mites) of denaturing solution; and (2) after precipitation with isopropanol and washing with 70% ethanol, the total RNA pellet was air-dried for 5 minutes and resuspended in 10 µL of diethyl pyrocarbonate water. The 10-µL sample of total RNA was used for the oligo dT–primed first cDNA synthesis using Thermoscript reverse transcriptase (GIBCO Life Technologies, Melbourne, Australia) with the second strand synthesized by use of *Pfu* polymerase and the following gene-specific oligonucleotide primers: Der p 1F, 5'-CTGGATCCCGTCCAT-

CATCGATCAAA-3'; Der p 1R, 5'-CTGGATCCTTTAGAGAAT-GACAACATATGGAT; Der p 2F, 5'-GCGGATCCTGATCAAGTC-GATGTCA-3'; and Der p 2R, 5'-TGGGATCCTTGTTTGATT-TAATCGC-3'.

The PCR fragments were *Bam*HI digested and cloned into *Bam*HI-digested pUC18 for sequencing. Sequencing was performed with an ABI50 sequencer (Perkin-Elmer Applied Biosystems, Melbourne, Australia) and the data analyzed using MACVector version 6.5.1 (Oxford Molecular Group, Towson, Md).

T-cell responses

The T-cell responses were measured as described previously.¹¹ In brief, PBMCs from 28 subjects (15 mite allergic and 13 nonallergic as determined by RAST) were stimulated with a set of 19-mer peptides progressing across Der p 1 by 5 residues to incorporate a 14-residue overlap. Peptides that contained positions 50, 124, and 136 were synthesized with both of the allelic polymorphisms. Proliferation was measured by thymidine incorporation assay. For 7 of the subjects, 41 peptides spanning the entire sequence were examined; for logistical reasons the remaining 21 subjects were stimulated with 30 peptides shown in a previous study to be stimulatory. For Der p 2, 21 subjects (allergic and nonallergic) were stimulated with 30 µg/mL of recombinant variants produced in *Escherichia coli* and purified as described by Mueller et al.¹² The dose was determined to be optimal in pilot experiments.

RESULTS

Wild mite reverse transcriptase PCR

The first batch of mites to be amplified were the dried mite cultures from Sydney. For the remaining mites, a reverse transcriptase PCR method was used. RNA was

			Der p 2 polymorphisms						
			40	47	98	111	114	116	127
<i>Commercial cDNA</i>	2.0101	(1)	V	T	A	M	D	V	I
	2.0102	(2)	V	S	A	M	N	V	L
	2.0103	(3)	L	S	A	M	N	V	I
<i>Environmental genomic</i>	2.0104	(a)	L	S	A	L	N	V	I
	2.0105	(b)	L	S	A	M	D	V	I
	2.0101	(c)	V	T	A	M	D	V	I
<i>Perth cDNA</i>	2.0105	(d)	L	S	A	M	D	V	I
	2.0106	(1)	V	T	T	M	D	V	I
	2.0104	(2)	L	S	A	L	N	V	I
	2.0107	(3)	V	S	A	M	D	V	I
	2.0107	(4)	V	S	A	M	D	V	I
	2.0108	(5)	V	S	A	M	D	A	I
	2.0101	(6)	V	T	A	M	D	V	I
	2.0101	(7)	V	T	A	M	D	V	I
	2.0101	(8)	V	T	A	M	D	V	I
	2.0101	(9)	V	T	A	M	D	V	I

FIG 1. Alignment of Der p 2 environmental and commercial sequence polymorphisms. Numbers at the top represent amino acid numbering based on the reference sequence Der p 2.0101.¹⁸ The polymorphic residues are highlighted in *red*. The CSL cDNA and genomic DNA clones were described previously by Chua et al.⁷ The correct nomenclature has been applied to the corresponding clones as indicated, with the original numbering in *parentheses*.

successfully extracted from every sample of mites, even when only a single mite was used as starting material. The RNA extracted from a single mite was sufficient for the synthesis of good yields of first-strand cDNA with gene-specific primers.

Der p 2 wild mite cDNA sequences

Nine Der p 2 cDNAs were isolated from the Perth wild mites. The sequences revealed 16 nucleotide changes in 15 positions, 6 of which resulted in an amino acid substitution. Table I shows the polymorphisms detected in this and other studies, both commercial and environmental. Only 1 of 9 sequences (clone 9) was identical to the published Der p 2.0101 nucleotide sequence from the commercial (CSL Ltd) mites. However, the sequences from clones 7 and 8 differed by only a single silent substitution at nucleotide 147, and clone 6 had an identical amino acid sequence but with 2 more silent changes at nucleotides 243 and 324. Each of these clones (6, 7, 8, and 9) came from a different house. Clone 1 had only a single nucleotide difference that resulted in a conservative alanine-to-threonine substitution at amino acid position 98. Clones 3 and 4 were isolated from different houses and were identical to each other, with several nucleotide changes resulting in silent changes (as described below) but with only a single amino acid substitution at amino acid position 47 of a threonine-to-serine residue.

The most consistent variation that resulted in an amino acid change was in the codon for amino acid residue 47, which encoded a serine in half of all the Der p 2 sequences and a threonine in the other half (Fig 1). The clones that contained a threonine at this position were the Der p 2.0101-like sequences described above and comprised the most common single sequence found. Those

clones that contained a serine at this position (Perth clones 2, 3, 4, and 5) showed a distinctive and related pattern of nucleotide variation. All had silent substitutions at nucleotides 168, 225, 234, and 324, indicative of a common sequence divergence. As shown for comparison in Table I, these silent changes were also found in previously published commercial (CSL Ltd) cDNA and environmental genomic DNA sequences from Sydney and Taipei, but again only in clones that contained a serine at position 47. This divergent sequence is thus widespread. Five of 9 of the serine 47 clones also showed characteristic changes resulting in amino acid substitutions at either 1 or several of the following residues: 40 (Val to Leu), 111 (Met to Leu), and 114 (Asp to Asn). In particular, Perth clone 2, genomic clone 1 (from Taipei), and CSL cDNA clone 3 contained all 4 substitutions.

In addition to the 4 amino acid polymorphisms above, changes at positions 98 Ala/Thr and 116 Val/Ala were found, but only in single clones. With respect to amino acid similarity grouping, most of these changes are conservative except the Asp/Gln substitution at position 114 and the rare Val/Ala change at position 116. The polymorphism described previously for residue 127 (in 1 commercial sequence) was not found in the Perth cDNA.

Der p 1 wild mite cDNA sequences

The 12 nucleotide sequences of Der p 1 cDNA from the Perth wild mites showed considerably less variation than the Der p 2 sequences (Table II). Of the 12 nucleotide substitutions found in the clones, 8 were restricted to a single clone; 3 were found in 2, 3, and 6 clones, respectively; and a single change at residue 491 was found in all 12 clones sequenced. None of the substitutions occurred together in a way that would suggest

TABLE II. Der p 1 environmental and commercial nucleotide polymorphisms

Nucleotide residue No.	Der p 1 polymorphisms			Clones containing polymorphisms			
	AA residue No.	Reference (Der p 1.0101) (triplet/AA)	Polymorphic (triplet/AA)	Environmental cDNA	CSL cDNA	Sydney genomic DNA	UK genomic DNA
298	-16*	GCA/Ala	GCT/Ala	3	—	ND	ND
398	19	ATG/Met	TTG/Leu	2	—	—	—
404	21	ACT/Thr	CCT/Pro	5	—	—	—
505	44	GAA/Glu	GAC/Asp	6	—	—	—
491	50	CAC/His	TAC/Tyr	All	c, d	1, 2, 3	1
553	70*	GGT/Gly	GGC/Gly	9, 10	—	—	—
584	81	GAA/Glu	AAA/Lys	—	d	—	—
714	124	GTA/Val	GCA/Ala	3, 4, 5, 6, 7, 9	a, e	3	1
714/715	124	GTA/Val	GCC/Ala	11	c, d	1, 2	—
717	125	AAC/Asn	AGC/Ser	7	—	—	—
728	129	GAA/Glu	AAA/Lys	8	—	—	—
749	136	AGC/Ser	ACC/Thr	—	e	3	—
757	138	ATT/Ile	ATG/Met	8	—	—	—
790	149*	GCA/Ala	GCT/Ala	1, 4, 10	a, c, d, e	—	—
825	161	CGC/Arg	CAC/His	—	—	1	—
986	215	GAA/Glu	AAA/Lys	7, 8	—	—	—
986	215	GAA/Glu	CAA/Gln	—	e	—	—

The nucleotide and corresponding amino acid numbering is based on the numbering for the full-length reference sequence of Der p 1.0101 (GenBank accession No. U11695).¹³ Environmental clones 1, 2, 3, and 10; clones 4, 5, and 9; and clones 6 and 7 were each isolated from the same houses. Clones 8 and 11 were isolated from 2 additional houses. The CSL cDNA was described previously by Chua et al,⁶ and the UK genomic clone was by Kent et al¹⁰ (revised sequence, accession No. X65197).

AA, Amino acid; ND, not determined.

*Silent substitutions.

a common pattern of inheritance. The substitution found in 6 of 12 clones at nucleotide 714 produced a change of amino acid from valine to alanine. Apart from this change, 7 of the Perth clones encoded an identical amino acid sequence, and 3 differed by 1 further residue unique to each clone. Clone 8 had 3 additional amino acid substitutions, and clone 7 had 2. Although compared with Der p 2 the clones had fewer nucleotide variations, most of the changes that did occur resulted in amino acid substitutions. Of the 11 nucleotide substitutions within the coding region of the mature protein, only 2 were silent. Of the 9 missense changes, 4 of the substitutions at residues 21, 125, 129, and 215, were nonconservative.

The substitutions found in the previously published CSL sequences are shown in Fig 2 for comparison. Except for the 124 Val/Ala exchange, the substitutions found in the wild Perth mites were not found in the CSL mites, and the CSL mites contained substitutions at positions 50 and 136 that were not found in the Perth sequences. A non-conservative change in residue 215 in both populations was found, but it was Glu to Gln in the CSL mites and Glu to Lys in the Perth mites. The exchange of Tyr and His at position 50 in the commercial mites is noteworthy because the sequence of Der p 1 that was first described contained His50,¹³ but all environmental sequences contain Tyr50. As with the Perth sequences and noted in Chua et al,⁶ most of the nucleotide substitutions in the CSL mites produced amino acid changes.

Der p 1 wild mite genomic DNA

Data on the genomic sequence from 3 mites from Sydney were obtained and are included in Table II and Fig 2

along with information from a sequence published for a mite cultured in England. These sequences all contain the Tyr50 and the 124 Ala/Val variation described for the Perth and CSL mites, with no silent changes. Two additional changes that resulted in conservative amino acid changes were found at positions 136 and 161, the change at position 136 also being found in a CSL sequence.

T-cell responses to Der p 1 peptide epitopes containing frequent polymorphisms

The responses of PBMCs to peptides of Der p 1 that incorporated both the frequent His/Tyr50 substitution found in all environmental sequences and the Ala/Val124 substitution that was found to be equally distributed in the sequences were investigated. The less frequent exchanges of Ser/Thr at position 136 and Glu/Gln at position 215 were also examined. PBMCs from 15 mite-allergic donors and 13 nonallergic donors were tested. As shown in other studies,^{11,14} no differences in the magnitude of responses of PBMCs from allergic and nonallergic subjects were evident.

The PBMCs of 19 of 28 (68%) subjects responded to at least 1 peptide. Of the 19 responders, 8 were stimulated by peptides that included amino acid positions 50, 124, and 136 (Table III), and 4 of these did not respond to any other peptides tested. Five of these subjects responded to peptide 110-128, and all of them had significant responses to only a single form of the peptide, 3 with valine and 2 with alanine. Two subjects responded to peptide 45-63 but only when the peptide contained a tyrosine at residue 50. One subject only responded to peptide 130-148 in the form containing the minor Thr136

		Der p 1 polymorphisms												
		19	21	44	50	81	124	125	129	136	138	161	215	
Commercial cDNA	1.0101 (a)	M	T	E	H	E	A	N	E	S	I	R	E	
	1.0102 (c)	M	T	E	Y	E	V	N	E	S	I	R	E	
	1.0103 (d)	M	T	E	Y	K	V	N	E	S	I	R	E	
	1.0104 (e)	M	T	E	H	E	A	N	E	T	I	R	Q	
UK genomic DNA	1.0105 (a)	M	T	E	Y	E	A	N	E	S	I	R	-	
Sydney genomic DNA	1.0106 (1)	M	T	E	Y	E	V	N	E	S	I	H	E	
	1.0102 (2)	M	T	E	Y	E	V	N	E	S	I	R	E	
Perth cDNA	1.0107 (3)	M	T	E	Y	E	A	N	E	T	I	-	-	
	1.0102 (1)	M	T	E	Y	E	V	N	E	S	I	R	E	
	1.0108 (2)	L	T	E	Y	E	V	N	E	S	I	R	E	
	1.0105 (3)	M	T	E	Y	E	A	N	E	S	I	R	E	
	1.0105 (4)	M	T	E	Y	E	A	N	E	S	I	R	E	
	1.0109 (5)	M	P	E	Y	E	A	N	E	S	I	R	E	
	1.0110 (6)	M	T	D	Y	E	A	N	E	S	I	R	E	
	1.0111 (7)	M	T	E	Y	E	A	S	E	S	I	R	K	
	1.0112 (8)	M	T	E	Y	E	V	N	K	S	M	R	K	
	1.0105 (9)	M	T	E	Y	E	A	N	E	S	I	R	E	
	1.0102 (10)	M	T	E	Y	E	V	N	E	S	I	R	E	
	1.0105 (11)	M	T	E	Y	E	A	N	E	S	I	R	E	
1.0102 (12)	M	T	E	Y	E	V	N	E	S	I	R	E		

FIG 2. Alignment of Der p 1 environmental and commercial sequence polymorphisms. The numbers at the top represent amino acid numbering based on the reference sequence Der p 1.0101 (accession No. U11695).¹³ The polymorphic residues are highlighted in *red*. The CSL cDNA was described previously by Chua et al,⁶ and the UK genomic clone was described by Kent et al¹⁰ (revised sequence, accession No. X65197). The correct nomenclature has been applied to the corresponding clones as indicated, with the original numbering system in *parentheses*.

variant. The sizes of the stimulation indices (3.3-8.1) were similar to those found for peptides without polymorphisms, with the exception of subject X, who had stimulation indices of 17 and 58 for 2 peptides that did not contain polymorphisms. None of the subjects responded to both forms of the peptides. Only 2 of 7 subjects (subjects K and L) who were tested with the full range of 41 peptides responded to a peptide that contained a polymorphism, with 1 of these subjects (subject L) responding to a single peptide only, namely the polymorphic peptide 110-128V. The responses of 3 of 21 of the remaining subjects (subjects O, Q, and Y) were confined to the polymorphic peptides, 45-63Y, 110-128A, and 136-148T; however, only 30 of the 41 peptides that span the Der p 1 sequence were tested for this group.

T-cell responses to variants of Der p 2

Der p 2 contains the highest density of polymorphic residues and can readily be produced in *E coli*. Recombinant proteins representing 4 of the variants—Der p 2.0101, 2.0104, 2.0107, and 2.0108—were produced, and their ability to stimulate PBMCs in in vitro proliferation assays was measured in 21 subjects (Fig 3). The alignment of the stimulation indices against the reference of Der p 2.0101 shows there was little variation in responses to the different polymorphisms. One subject did respond to Der p 2.0107 with a higher stimulation index than the other variants, but the response to the other variants was already high. The correlation coefficients were 0.94, 0.73, and 0.89 for the comparison of Der p 2.0101 with Der p 2.0104, 2.0107, and 2.0108. The Der p 2.0104 molecule with the highest sequence divergence had the best correlation.

TABLE III. Proliferative responses of PBMC to peptides of Der p 1 containing polymorphic residues

Subject	50 H/Y	124A/V	136S/T	215E/Q	Other
K*	—	110-128V (3.3)	—	—	105-123 (6.9)
L	—	110-128V (5.1)	—	—	—
O	45-63Y (8.1)	—	—	200-218Q (3.9)	—
Q*	45-63Y (4.0)	—	—	—	—
T*	—	110-128A (3.4)	—	—	105-123 (3.1)
U	—	110-128V (3.7)	—	—	105-123 (7.8)
X*	—	—	130-148T (3.4)	—	105-123 (58.2), 80-98 (17.7)
Y	—	110-128A (4.2)	—	—	—

This table shows the peptides used in the assay and the stimulation indices of these peptides only for responses with stimulation indices greater than 3 and a significance level of $P < 0.05$ by the Student *t* test. Only 8 of the 28 subjects responding to peptides that contain polymorphisms are shown. Donors K and L were stimulated with all 41 peptides that span the entire molecule, whereas donors O, Q, T, U, X, and Y were stimulated with 30 peptides that had been found to be stimulatory on previous occasions.

*Subject with mite allergy.

DISCUSSION

Der p 1 and Der p 2 are major allergic specificities and targets for new immunotherapeutic strategies such as the injection of peptides, modified allergens, or DNA vac-

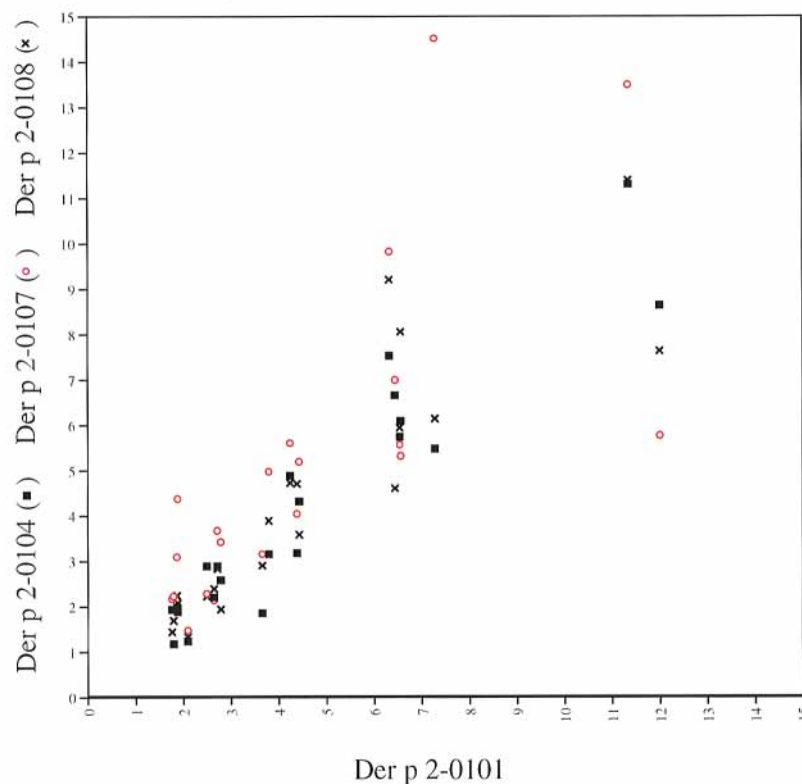


FIG 3. Proliferative responses of PBMCs to polymorphic variants of Der p 2. PBMCs from 21 subjects were stimulated with recombinant variants of Der p 2.0101, 2.0104, 2.0107, and 2.0108. The mean stimulation indices of the responses of each individual to Der p 2.0104, 2.0107, and 2.0108 are shown plotted against their response to Der p 2.0101 as a reference. The correlation coefficients versus Der p 2.0101 were 0.94, 0.73, and 0.89, respectively, for 2.0104, 2.0107, and 2.0108.

cines. cDNA has provided the information to pursue these objectives but has also revealed that they exhibit a high frequency of allelic polymorphism. This study examined which variants found in the environment could best represent the spectrum of polymorphic proteins that induce sensitization. It was found that although the sequences of cDNA clones from wild mites did vary, the variation was within a defined pattern that can be represented by 1 or 2 sequences already determined from mite cultures maintained in commercial production. The pattern of genetic variation determined for Der p 1 and Der p 2 was different.

The sequences of the Der p 1 cDNA clones frequently contained nucleotide substitutions that result in amino acid substitutions. The nucleotide sequences showed far less variation than the Der p 2 sequences, with most changes resulting in a missense substitution and only a few silent substitutions. This was previously noted for the sequences from the commercial CSL cultures. The substitutions are unlikely to be artifacts of the cloning because the same substitutions were found in independent cDNA preparations (eg, amino acid 215 in clones 7 and 8 and amino acid 136 in the CSL and Sydney sequences). Moreover, 8 clones of Der p 4 and Der p 5, both environmental and commercial, and the Der p 2 clones described in this study were all examined with the

same techniques and have not shown this kind of variation. All but 2 of the substitutions in Der p 1 were only found in 1 or 2 of the 20 sequences presented here. The first exception was at position 124, which had an equal distribution of sequences encoding either an alanine or a valine. Because the changes at other positions were rare, the variants designated Der p 1.0102 and Der p 1.0105, differing by a single valine/alanine exchange, would be highly representative of the dominant sequences. The second exception is at residue 50, involving a histidine/tyrosine substitution. It should be noted that the first published sequence of Der p 1.0101 from a CSL commercial culture differs from all environmental isolates in that it has a histidine substitution at position 50, whereas all the environmental isolates have a tyrosine.¹³ The only other sequence with the histidine at residue 50 was another cDNA from the commercial mites.

The group 1 molecules, which have 2 globular domains, show 82% identity between the species, with 50% of the amino acid substitutions being nonconservative. The divergence between Der f 1 and Der p 1 is found predominantly in regions 1-20, 201-222, and 91-130.⁶ The allelic polymorphisms are spread throughout the structure of the Der p 1. According to the model of Topham, as described by Robinson et al,¹⁵ only the rare

substitutions at positions 136 and 138 are in solvent-inaccessible positions and are therefore likely to affect structure. However, even conservative changes of key determinants of stability or activity can have severe structural or phenotypical effects.¹⁶ The high frequency of missense substitutions suggests that some Der p 1 substitutions can confer a selective advantage, but they cannot be maintained in the absence of the normal enzyme activity. For example, a change in protein cleavage specificity may help the processing of food, but only if the original enzyme function is also preserved. Der p 1 has been shown to have a mixed serine and cysteine protease specificity, so it would be interesting to examine the specificity of these polymorphic proteins.¹⁷

Der p 2 polymorphisms were essentially limited to 4 previously described amino acid substitutions, with the pattern of substitution highlighting an evolutionary divergence of 2 sequences. The most abundant sequence was the first described, Der p 2.0101.¹⁸ It was found in 4 of 9 sequences determined from the Perth environmental mites. An additional 3 sequences differed by only a single amino acid substitution. Several Der p 2 clones, including those from the Perth mites and mites from other environmental and commercial sources, showed a distinct evolution toward a second sequence represented by Der p 2.0104. Der p 2.0104 has 4 amino acid substitutions (LSLN at positions 40, 47, 111, and 117) when compared with Der p 2.0101. All but 3 of the amino acid polymorphisms in the 16 sequences were variations on this pattern. Like Der p 1, one amino acid position, residue 47, showed an equal distribution of a substitution, either serine or threonine. The sequences with serine showed the silent and missense changes exemplified by Der p 2.0104. The silent changes at nucleotides 56, 75, 78, and 108 were conserved in all of the Der p 2.0104-like clones. The presence of the same silent substitutions in mites from different sources shows a limited common origin of genetic variation. The reason for the sequence diversity in the Der p 2 sequences might be that if Der p 2 protects mites from bacterial pathogens,¹⁹ a reservoir of mites with Der p 2 variants with different binding activities to different pathogens may help protect the species from zoonoses.

The single structural domain of the 14-kd group 2 allergens has been noted for its high sequence identity (88%) among *D pteronyssinus*, *D farinae*, and *Euroglyphus maynei*,²⁰ with the changes having little predicted effect on protein structure.²¹ Similarly, the polymorphisms are also largely conservative. The substitutions of 40, 47, 111, and 114 that tend to occur together lie in close proximity within the single domain structure of Der p 2. Residues 47 and 114 lie in adjacent loop structures, residue 111 is at the termini of an adjacent β strand, and residue 40 is within the 35-44 β strand.²¹ It is most likely that these substitutions would have little effect on the overall tertiary structure. With respect to T-cell epitopes, O'Hehir et al²² showed that T cells from 7 of 18 subjects recognized peptides representing residues 22-50 and 105-129, regions affected by polymorphisms. However, peptides encompassing regions 61-86 and 78-104 that do not

include polymorphic residues were more stimulatory (each stimulating 12/18 samples), and no subject responded only to peptides that encompassed polymorphic sites. A similar study of O'Brien et al²³ found that the most frequent response was to peptide 111-129, which induced proliferation in 16 of 24 samples. From the latter study, it could be speculated that the substitutions of residues 111 and 114 could have a significant influence on the T-cell responses. For both studies it should be noted that only the responses to the Der p 2.0101 sequences were examined, and hence any responses dependent on the substituted residues were not determined.

The T cells of some individuals may respond to rare epitopes produced by sporadic amino acid substitutions, but because individuals recognize more than 1 epitope on an allergen, most subjects should respond to the conserved sequences. T cells from persons living in a region with its acarofauna dominated by *D pteronyssinus* have been shown to produce similar proliferative and cytokine responses to Der p 1 and Der f 1 despite an 18% sequence divergence. This was also the case for the responses to the 4 variants of Der p 2, which differed from each other by up to 4 residues, including the closely spaced 40 and 47, as well as 111, 116, and 127. All 21 subjects examined had similar responses to each variant, except for a high responder who had an even higher response to 1 variant. In contrast, experiments with peptides from different species have indicated a more stringent sequence requirement, and similar results were obtained here. The PBMCs from 5 of 28 subjects responded to peptide 110-128 containing the frequent polymorphism at position 124, with each subject preferentially recognizing 1 allelic form. The PBMCs from 2 of 28 subjects recognized peptide 45-63 and were only stimulated by the tyrosine 50 form, typical of the environmental isolates. Thus immunotherapeutic strategies with peptides may need to consider even single sequence substitutions for the most effective outcome. Although their data are limited, of the 5 synthetic peptides reported by others to stimulate T-cells clones, 45-73, 89-117, 111-139,²⁴ 110-119, and 110-131,²⁵ 3 contain polymorphic residues. The polymorphisms may play a role in allergenicity by providing different peptides that can be presented by a wider variety of MHC haplotypes. For example, the allergens Der p 4 and Der p 5, with only single amino acid substitutions, induce high-titer IgE responses in only 50% of mite-allergic subjects, whereas the group 1 and 2 allergens induce IgE in approximately 90%.

The sequence variation of Der p 1 and Der p 2 found in the wild mites was essentially limited to that already described from the mites maintained in culture by CSL Ltd. Despite the microheterogeneity, 2 sequences that were highly representative of the genetic variation in mites could be identified for each allergen. The data from the mites from Perth, Sydney, Taipei, and the UK, as well as the mites from CSL Ltd, all fit the same pattern of genetic variation. This includes the silent substitutions of Der p 2 that would not be influenced by natural selective pressure. It will be of interest to ascertain whether sequences from other regions follow this pattern.

Finally, an important technical advance produced by this study was to use a small-scale reverse transcriptase PCR that could conveniently be used to synthesize cDNA from a single mite. The need to culture mites to produce sufficient starting material for cDNA synthesis, for further genetic analysis or to produce recombinant allergens, has been circumvented.

We would like to thank Dr Euan Tovey, from the Institute of Respiratory Medicine, University of Sydney, Australia, for collecting and preparing the dried house dust mite samples from the Sydney homes.

REFERENCES

- Rafnar T, Griffith IJ, Kuo MC, Bond JF, Rogers BL, Klapper DG. Cloning of Amb a I (antigen E), the major allergen family of short ragweed pollen. *J Biol Chem* 1991;266:1229-36.
- Ong EK, Griffith IJ, Knox RB, Singh MB. Cloning of a cDNA encoding a group-V (group-IX) allergen isoform from rye-grass pollen that demonstrates specific antigenic immunoreactivity. *Gene* 1993;134:235-40.
- Suphioglu C, Singh MB. Cloning, sequencing and expression in *Escherichia coli* of Pha a 1 and four isoforms of Pha a 5, the major allergens of canary grass pollen. *Clin Exp Allergy* 1995;25:853-65.
- Ferreira F, Hirtenlehner K, Jilek A, et al. Dissection of immunoglobulin E and T lymphocyte reactivity of isoforms of the major birch pollen allergen Bet v 1: potential use of hypoallergenic isoforms for immunotherapy. *J Exp Med* 1996;183:599-609.
- Griffith IJ, Craig S, Pollock J, Yu XB, Morgenstern JP, Rogers BL. Expression and genomic structure of the genes encoding Fel d I, the major allergen from the domestic cat. *Gene* 1992;113:263-8.
- Chua KY, Kehal PK, Thomas WR. Sequence polymorphisms of cDNA clones encoding the mite allergen Der p I. *Int Arch Allergy Immunol* 1993;101:364-8.
- Chua KY, Huang CH, Shen HD, Thomas WR. Analysis of sequence polymorphism of a major mite allergen, Der p 2. *Clin Exp Allergy* 1996;26:829-37.
- Nishiyama C, Yuuki T, Takai T, Okumura Y, Okudaira H. Determination of three disulfide bonds in a major house dust mite allergen, Der f II. *Int Arch Allergy Immunol* 1993;101:159-66.
- Olsson S, van Hage-Hamsten M, Whitley P, et al. Expression of two isoforms of Lep d 2, the major allergen of *Lepidoglyphus destructor*, in both prokaryotic and eukaryotic systems. *Clin Exp Allergy* 1998;28:984-91.
- Kent NA, Hill MR, Keen JN, Holland PWH, Hart BJ. Molecular characterization of the group I allergen Eur m I from house dust mite *Euroglyphus maynei*. *Int Arch Allergy Immunol* 1992;99:150-2.
- Hales BJ, Thomas WR. T-cell sensitization to epitopes from the house dust mites *Dermatophagoides pteronyssinus* and *Euroglyphus maynei*. *Clin Exp Allergy* 1997;27:868-75.
- Mueller GA, Smith AM, Williams DC Jr, et al. Expression and secondary structure determination by NMR methods of the major house dust mite allergen Der p 2. *J Biol Chem* 1997;272:26893-8.
- Chua KY, Stewart GA, Thomas WR, et al. Sequence analysis of cDNA coding for a major house dust mite allergen, Der p I. Homology with cysteine proteases. *J Exp Med* 1988;167:175-82.
- Hales BJ, Shen H-D, Thomas WR. Cross-reactivity of T-cell responses to *Dermatophagoides pteronyssinus* and *D. farinae*. Studies with group 1 and 7 allergens. *Clin Exp Allergy* 2000;30:927-33.
- Robinson C, Kalsheker NA, Srinivasan N, et al. On the potential significance of the enzymatic activity of mite allergens to immunogenicity. Clues to structure and function revealed by molecular characterization. *Clin Exp Allergy* 1997;27:10-21.
- Pakula AA, Sauer RT. Genetic analysis of protein stability and function. *Annu Rev Genet* 1989;23:289-310.
- Hewitt CR, Horton H, Jones RM, Pritchard DI. Heterogeneous proteolytic specificity and activity of the house dust mite proteinase allergen Der p 1. *Clin Exp Allergy* 1997;27:201-7.
- Chua KY, Doyle CR, Simpson RJ, Turner KJ, Stewart GA, Thomas WR. Isolation of cDNA coding for the major mite allergen Der pII by IgE plaque immunoassay. *Int Arch Allergy Appl Immunol* 1990;91:118-23.
- Ichikawa S, Hatanaka H, Yuuki T, et al. Solution structure of Der f 2, the major mite allergen for atopic diseases. *J Biol Chem* 1998;273:356-60.
- Smith W, Mills K, Hazell L, Hart B, Thomas W. Molecular analysis of the group 1 and 2 allergens from the house dust mite, *Euroglyphus maynei*. *Int Arch Allergy Immunol* 1999;118:15-22.
- Mueller GA, Benjamin DC, Rule GS. Tertiary structure of the major house dust mite allergen Der p 2: sequential and structural homologies. *Biochemistry* 1998;37:12707-14.
- O'Hehir RE, Verhoef A, Panagiotopoulou E, et al. Analysis of human T cell responses to the group II allergen of *Dermatophagoides* species: localization of major antigenic sites. *J Allergy Clin Immunol* 1993;92:105-13.
- O'Brien RM, Thomas WR, Nicholson I, Lamb JR, Tait BD. An immunogenetic analysis of the T-cell recognition of the major house dust mite allergen Der p 2: identification of high- and low-responder HLD-DQ alleles and localization of T-cell epitopes. *Immunology* 1995;86:176-82.
- O'Hehir RE, Garman RD, Greenstein JL, Lamb JR. The specificity and regulation of T-cell responsiveness to allergens. *Annu Rev Immunol* 1991;9:67-95.
- Higgins JA, Thorpe CJ, Hayball JD, O'Hehir RE, Lamb JR. Overlapping T-cell epitopes in the group I allergen of *Dermatophagoides* species restricted by HLA-DP and HLA-DR class II molecules. *J Allergy Clin Immunol* 1994;93:891-9.