

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

C. SUPHIOGLU and M. B. SINGH

School of Botany, The University of Melbourne, Parkville, Victoria, Australia

Summary

Background The pollen of canary grass, which was introduced as a pasture grass from Europe, is a major allergen source in the external environment of southern Australia. This study was performed to characterize the major recombinant allergens of canary grass pollen. It is anticipated that recombinant allergens may be useful in diagnosis and immunotherapy of grass pollen induced allergies.

Objective To clone major canary grass pollen allergens and assess their nucleotide and amino acid sequence homologies with other grass pollen allergens. This sequence information may then be useful in T and B cell epitope mapping studies.

Methods A canary grass pollen λ gt11 cDNA expression library was constructed and screened with sera of grass-pollen-sensitive patients. IgE-reactive clones were isolated, sub-cloned into *Escherichia coli*, sequenced and, along with the deduced amino acid sequences, compared with other sequences in nucleotide and amino acid databases.

Results One of the clones encoded the group 1 allergen of canary grass pollen, Pha a 1, with a deduced amino acid sequence identity of 88.8% with Lol p 1, from rye-grass pollen, 88.1% with Hol 1 1, from velvet grass pollen and 86.6% with Phl p 1, from timothy grass pollen. The other clones (e.g. clones, 5, 14, 28, 29) encoded polymorphic forms of Pha a 5. These polymorphic forms showed between 60.6–95.5% nucleotide and 40.1–81.7% deduced amino acid sequence identities with each other. Moreover, they shared significant sequence identity with other group 5 allergens from rye-grass, timothy and Kentucky bluegrass pollens.

Conclusions Group 1 and four isoforms of group 5 allergens of canary grass pollen have been cloned and upon sequencing demonstrated strong nucleotide and amino acid sequence identities with other group 1 and 5 grass pollen allergens.

Keywords: Canary grass, *Phalaris aquatica*, Pha a 1, Pha a 5, allergen, cDNA cloning, isoallergens, nucleotide sequence, amino acid sequence

Clinical and Experimental Allergy, Vol. 25, pp. 853–865. Submitted 15 August 1994; revised 7 February 1995; accepted 8 April 1995.

Introduction

Substantial proportions (15–20%) of the populations in the developed countries suffer from IgE-mediated atopic disease such as hay fever and bronchial asthma [1–2]. Canary grass (*Phalaris aquatica*) pollens represent major

sources of allergenic proteins in southern Australia, especially in Victoria [3]. The isolation and characterization of several protein allergens from canary grass pollens have been recently reported [4]. One of the major allergens from canary grass pollen, with a molecular mass of 34 kDa, has been isolated and its NH₂-terminal sequence has been determined [4]. The sequence of this protein showed $\geq 90\%$ identity to that reported for rye grass allergen Lol p 1, and other group 1 allergens

Correspondence: Dr M. B. Singh, School of Botany, The University of Melbourne, Parkville, Victoria, 3052, Australia.

from timothy, and Kentucky bluegrass pollens [5–7], and was thus tentatively designated Pha a 1. Moreover, group 5 allergens were also present in canary grass pollen. However, the definitive designation for Pha a 1 and Pha a 5 requires the full sequence characterization in order to assess their degree of overall sequence identity with other group 1 and group 5 allergens. This has now been established and is reported in this paper.

Analysis of cDNA-encoding allergens has indeed already made a significant contribution to defining allergens and providing material and information for epitope and T cell studies. It is anticipated that its usefulness could extend into developing reagents to improve diagnosis and immunotherapy. Several allergens, including Amb a 1, 2, 5, 6 from ragweed pollen [8–10], Poa p 1 [11] and Poa p 5 (previously known as Poa p 9) from Kentucky bluegrass pollen [12], Lol p 1 [6,13,14] and Lol p 5 [15,16] from rye-grass pollen and Der p 1 from house dust mite [17], have been shown to be polymorphic by either protein sequence analysis or comparison of predicted protein sequence of isolated cDNA clones. In the case of polymorphic forms of Lol p 5, Lol p 5A [15] has 66.4% identity and 80.4% similarity with Lol p 5B [16]. Both Lol p 5A and Lol p 5B possess unique as well as crossreactive antigenic determinants [16]. These variants exhibit slight physicochemical differences but possess immunologically closely related properties. The existence of these sequences has highlighted the importance of determining the polymorphism in the allergens in defining which isoforms in a family of related proteins play a role in allergenic properties.

This paper presents the cloning, expression and sequencing of cDNA clones encoding Pha a 1 and polymorphic forms of Pha a 5 allergens from canary grass pollen. Nucleotide and deduced amino acid sequences of clones encoding Pha a 1 and polymorphic forms of Pha a 5 are shown to possess high homologies with group 1 and group 5 allergens, respectively, of other clinically important grass pollens.

Materials and methods

Pollen

Pollen was collected from canary grass (*Phalaris aquatica*) flowering heads in the field, on hot and still mornings during the flowering season in spring, and stored under liquid nitrogen until use. These pollen grains were characterized for morphology and purity by scanning electron microscopy according to the method of Staff *et al.* [18], in a JEOL JSM-840 scanning electron microscope at an accelerating voltage of 5 kV.

Sera

Sera were collected from five subjects with a previous clinical record of spring hay fever symptoms, who had not earlier undergone hyposensitization treatment, positive skin-prick test and radioallergosorbent test (RAST) towards crude rye-grass pollen extract. The RAST scores for all sera was four. These sera were also selected on the basis of high IgE-reactivity to crude canary grass pollen extract [4]. All sera were stored at -20°C in 0.5–1 mL aliquots. RAST was performed according to the manufacturer's instructions (Phadezyme RAST, Pharmacia LKB, Uppsala, Sweden).

Monoclonal antibodies

Monoclonal antibody (MoAb) FMC A1 (40.1) and FMC A7 (12.3) were produced in 1983 [19] by immunization of mice with crude rye-grass pollen extract and screening for MoAbs which recognized IgE binding proteins. The specificity of MoAb A1 was ascribed to a 35 kDa protein now known as Lol p 1 while MoAb A7 was found specific to a 28/30 kDa protein now known as Lol p 5. Monoclonal antibody 3A2 was produced by immunization of mice with purified Cyn d 1 allergen from Bermuda grass pollen [20].

Isolation of RNA

RNA was isolated using a modification of the guanidinium thiocyanate method of Chomczynski and Sacchi [21]. Briefly, canary grass pollen was ground with a mortar and pestle in liquid N_2 , with 9 mL of guanidinium thiocyanate buffer (5 M guanidinium thiocyanate in 0.05% w/v tris-HCl [pH 7.0], 0.05 volume of β -mercaptoethanol, 0.1 volume of 5% w/v sarcosyl). An equal volume of phenol was added to the pollen solution and it was shaken for 10 min, after which 1/2 volume of chloroform: isoamyl alcohol (24:1, v/v) was added. The mixture was shaken for a further 20 min followed by centrifugation at 7000 g for 25 min. The aqueous phase was collected and re-extracted with phenol:chloroform:isoamyl alcohol (25:24:1) followed by centrifugation at 2000 g until the interface was clear.

The aqueous phase was then decanted into a quickseal ultracentrifuge tube on top of a 3 mL CsCl cushion (5.7 M CsCl in 0.1 M EDTA; density = 1.71 g/mL) and centrifuged for 20 h at 40 000 rpm and 20°C in a Ti 70.1 rotor (Beckman L8-70 ultracentrifuge). After centrifugation the RNA pellet was resuspended in 0.05% w/v SDS, extracted with phenol/chloroform and precipitated overnight at -70°C in the presence of sodium acetate and the

RNA pellet was dissolved in 10 mM Tris-HCl [pH 7.2], 1 mM EDTA (TE buffer). The concentration and purity of the RNA was measured by determination of the optical density of the solution at 260 and 280 nm and comparison with a known standard.

Isolation of mRNA

Poly (A)⁺ mRNA was isolated using a mRNA purification kit following the manufacturers instructions (Pharmacia-LKB, Uppsala, Sweden). The concentration and purity of the mRNA was measured by determination of the optical density of the solution at 260 and 280 nm and comparison with a known standard.

Construction of cDNA libraries

A cDNA library was constructed from poly (A)⁺ mRNA, using a cDNA synthesis kit according to the manufacturers instructions (Pharmacia-LKB, Uppsala, Sweden). Linkered cDNA was inserted into the expression vector λ gt11 (Promega, WI, USA).

Immunological screening of cDNA libraries

Immunological screening was accomplished by incubating the libraries, or positive λ gt11 clones, with *E. coli* Y1090 cells at 37°C for 15 min and then plating onto Luria broth (LB) plates containing 100 mg/mL ampicillin (LA-plates). After 3 h growth at 42°C, the plates were overlaid with isopropyl- β -D-thiogalactopyranoside (IPTG) impregnated nitrocellulose (NC) filters (Hybond-C extra, Amersham, UK) in order to induce production of fusion proteins. The plates were then incubated for 4 h at 37°C. Replica filters were obtained by overlaying the plate with a second filter for another 4 h.

The replica filters were screened with a pool of five highly grass pollen-sensitive patient sera (diluted 1:4 in phosphate buffered saline with 0.5% w/v bovine serum albumin (PBS/BSA), which was pre-incubated with 1 mg/mL of *E. coli* extract [Promega, WI, USA] for 1 h at room temperature). The bound antibodies, on one of the filters, were detected using rabbit immunoglobulins to human IgE (Dakopatts, CA, USA; diluted 1:200 in PBS/BSA) followed by incubation in peroxidase-conjugated anti-rabbit Ig (Promega, WI, USA; diluted 1:2500 in PBS/BSA) for the detection of human IgE antibodies. Colour was developed using standard chromogenic methods described elsewhere [4]. The bound IgE antibodies of the replica filter were detected using ¹²⁵I-anti-human IgE (Kallestad, TX, USA; diluted 1:7 in PBS with 0.5% w/v BSA) as described elsewhere [4]. Positive

clones common with both detection methods were identified, single positive plaques cut from the plates and stored in SM buffer (50 mM Tris-HCl, pH 7.4, 100 mM NaCl, 10 mM MgSO₄, 0.1% w/v gelatin) at 4°C for at least 12 h before plating.

Dot-blot analysis of fusion proteins from λ gt11 clones

The λ gt11 clones encoding human IgE-reactive recombinant polypeptides of canary grass pollen were plated on Y1090 bacterial lawns and allowed to grow overnight at 37°C. Single plaques were transferred to 20 μ L of SM buffer with plastic pipette tips and allowed to sit for 1 h; 2 μ L of each clone was dotted on freshly plated Y1090 bacterial lawns and allowed to grow for 3–4 h at 42°C. Fusion-protein plaque lifts were then taken by overlaying IPTG-impregnated NC filters for 3–4 h, blocked with 10% w/v milk powder in PBS and the bound human IgE were then detected using standard chromogenic methods described earlier.

Isolation of λ DNA

Lambda-lysates were made by incubating an isolated plaque in 500 μ L of plating cells for 15 min, after which the volume was made up to 5 mL by addition of LB containing 100 mg/mL ampicillin, 5 mM CaCl₂, 0.4% w/v maltose and the culture incubated at 37°C, with shaking, for 5 h or until the bacterial cells were lysed; thereafter 500 μ L of chloroform were added to each culture and they were shaken for a further 15 min. Bacterial debris was removed by centrifugation at 8000 g for 10 min. The lysate was stored in the dark at 4°C until required. λ DNA was isolated from liquid lysates using the method of Messe *et al.* [22] or by Qiagen-tip 20 (Qiagen, Germany) according to the manufacturer's instructions. DNA was digested with restriction endonucleases and separated on agarose gels.

Subcloning of cDNA inserts into plasmid vectors

cDNA inserts from λ gt11 clones were isolated from agarose gels with Prep-A-Gene (Bio-Rad, CA, USA) and subcloned into pBluescript KS vectors (Stratagene, CA, USA) according to the manufacturers' instructions.

Oligonucleotide synthesis

Oligonucleotides were synthesized on Gene Assembler Plus (Pharmacia-LKB, Uppsala, Sweden) according to the manufacturers instructions. These oligonucleotides were derived from the internal partial sequences of the

cDNA clones and were used as sequencing primers in order to obtain complete sequence of the clones. Table 1 summarizes all the sequencing primers used in this study.

Nucleotide sequencing

For double-stranded DNA sequencing, template DNA was isolated using Qiagen-tip 20 (Qiagen, Germany) plasmid Mini preparation according to the manufacturer's instructions. DNA sequence was determined by double-stranded sequencing using the dideoxy chain termination method [23] with a T7 DNA sequencing kit (Pharmacia, Sweden) according to the manufacturer's instructions. In addition to the promoter primers, T3 and T7 (Promega, WI, USA), several specific oligonucleotides were used (for sequences see Table 1).

After sequencing reactions were completed, the samples were heated at 75–85°C for 2 min, loaded onto 6% w/v polyacrylamide/7 M urea gel (Sequagel, Kimberly Research, NJ, USA) and nucleotides were separated by electrophoresis at constant power of 50 W. The gel was transferred onto Whatman 3 mm paper, dried at 80°C under vacuum and exposed to X-ray film (XAR, Kodak) at room temperature.

Sequence analysis

Computer programs contained in PC/GENE (IntelliGenetics, CA, USA), MacProMass 2 (Beckman Research Institute, CA, USA) and DNA Strider (Commissariat à l'Energie Atomique, France) were used for both deduced amino acid and nucleotide

sequence analysis. Sequence homology database searches were performed with the NCSA/BYU Telnet program (University of Illinois, Urbana, USA) on an Apple McIntosh computer (Apple Computer Inc., CA, USA).

Results

Immunological identification of recombinant proteins

A canary grass pollen λ gt11 cDNA expression library was screened with pooled sera of grass-pollen-allergic patients. Out of a total of 1×10^4 plaques screened, 31 clones that bound to IgE antibodies were identified and plaque-purified.

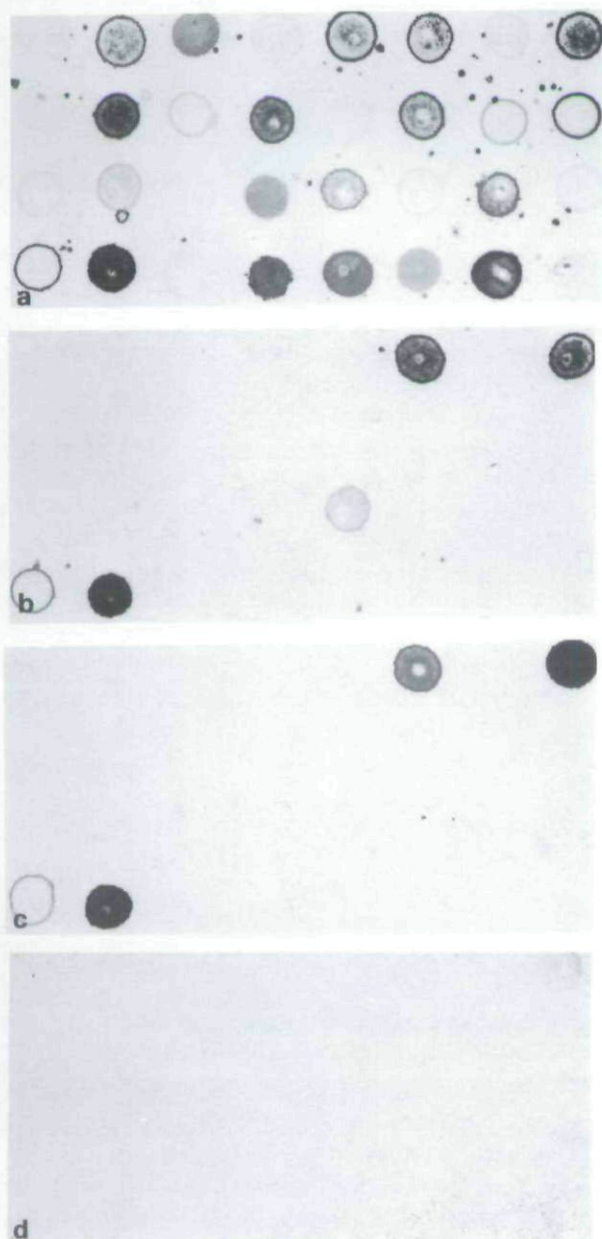
The purified clones gave no signal when tested with non-allergic human sera. Similarly, non-recombinant λ gt11 did not bind IgE antibodies. The ability to bind to specific IgE antibodies in sera of grass-pollen-sensitive patients identified these clones as encoding allergens.

The proteins expressed by cDNA of the 31 positive clones were tested for binding to group 1- and group 5-specific MoAbs as well as a serum pool of grass-pollen-allergic patients by dot blots (Fig. 1). Lol p 5A-specific MoAb A7 did not appear to bind to fusion proteins from any of the clones. Although it has been shown that MoAb A7 binds strongly to Lol p 5A [15], it does not bind to its isoform Lol p 5B [16]. Thus, some of these clones may encode group 5 proteins that may not possess the MoAb A7-binding epitope(s). However, group 1-specific MoAbs A1 and 3A2 bound strongly to fusion proteins of 4 cDNA clones (e.g. clones, 6, 8, 25

Table 1. Oligonucleotides used in sequencing Group 1 (GI) and Group 5 (GV) allergens of canary grass pollen. The oligomers match coding or non-coding strand sequences of clones as noted. Oligomers GV.1 and GV.3 are based on the semi-identical nucleotide sequences of clones 14, 28, 29 and clones 28, 29, respectively. Nucleotide position numbering for clone 6 corresponds to that used in Fig. 2

Name	Nucleotide sequence	Clone specificity	Strand	Nucleotide position
GI.1	5'-CCGTCGACGTACTIONTGA	6	Non-coding	589 ← 604
GI.2	5'-GTGCCGTCCGGGTACTT	6	Non-coding	526 ← 542
GI.3	5'-GAGTACGGCGACAAGTGGC	6	Coding	145 → 163
GI.4	5'-TCCATGGCCAAGAAGGGC	6	Coding	448 → 465
GV.1	5'-AGAATTAGTTATCAGCAAAA	14	Non-coding	1061 ← 1080
		28	Non-coding	1085 ← 1106
		29	Non-coding	1059 ← 1078
GV.2	5'-GATTAGACTTTGTAGCCGCC	5	Non-coding	511 ← 530
GV.3	5'-GAGCAGAAGCTGATCGAGGA	28	Coding	165 → 182
		29	Coding	196 → 215
GV.4	5'-CCGGCCACCCCTGCCGCCTC	14	Coding	145 → 166

C1	C2	C3	C4	C5	C6	C7	C8
C9	C10	C11	C12	C13	C14	C15	C16
C17	C18	C19	C20	C21	C22	C23	C24
C25	C26	C27	C28	C29	C30	C31	λ



and 26). Moreover, fusion proteins of 18 cDNA clones (e.g. clones, 2, 3, 5, 6, 8, 10, 12, 14, 16, 20, 21, 23, 25, 28, 29, 30 and 31) bound IgE antibodies very strongly. On the basis of strong MoAb and human IgE antibody binding, seven cDNA clones (e.g. clones 5, 6, 8, 14, 23, 28 and 29) were chosen for further study. Moreover, these seven clones, when tested with individual sera that made up the original serum pool, displayed very strong IgE-reactivity (data not shown).

Isolation and characterization of a cDNA clone encoding Pha a 1

Two cDNA clones (e.g. clones 6 and 8) that bound group 1-specific MoAbs A1 and 3A2 were subcloned into pBluescript plasmid vectors for sequencing. A range of oligonucleotides, based on the internal partial sequences of the two clones, were used to sequence the cDNA of clones 6 and 8 (Table 1). The nucleotide sequences of both clones were identical to each other. However, clone 8 did not include a poly (A)⁺ tail and thus the sequence of clone 6 was characterized further.

The nucleotide sequence of cDNA clone 6 is shown in Fig. 2. There is an open reading frame (ORF) of 807 base pairs (bp) starting with an ATG initiation codon at position 22 and terminating with a TGA stop codon at position 829. This ORF corresponds to a polypeptide of 269 amino acids with an estimated molecular mass of 29 kDa. The proposed translation initiation site and its flanking sequences (nucleotides 15–23) share 44% identity with the consensus plant sequence AACAAATGGC [24]. The most critical nucleotide, a purine at position –3, is conserved [24]. The cDNA has a complete 3'-untranslated region that contains the canonical AATAAA polyadenylation signal [25] followed by a poly (A)⁺ tail.

The predicted protein sequence has a hydrophobic putative signal peptide sequence of 29 amino acids. This is indicative of a mature processed protein of 240 amino acids with an estimated molecular mass of

Fig. 1. Specificity of IgE and MoAbs for clones expressing allergens of canary grass pollen. Recombinant polypeptides of clones 1 to 31 were expressed in λ gt11 and the fusion-proteins transferred to nitrocellulose. The fusion-proteins were screened for binding of IgE from serum pool of five grass-pollen-allergic individuals (a); Cyn d 1-specific MoAb 3A2 (b); Lol p 1-specific MoAb FMC A1 (c); Lol p 5A-specific MoAb FMC A7 (d). The position of the recombinant proteins produced by the different clones are shown in the template above. Proteins expressed by non-recombinant λ gt11 were used as negative controls (denoted by λ).

		-20		-10	
1	CAACAATTCT CAATTTTGAA	GATGATGAAG ATGGTGTGTT	CGTCGTCGTC GTCGTCGCTT	CTCGTGGTGG CGGCGCTGTT	80
		M M K M V C	S S S S S S L L V V	A A L L	
		1		10	
81	GGCCGTGTTT GTGGGCAGCG	CGCAGGGCAT CGCGAAGGTT	CCCCCTGGCC CGAACATCAC	GGCCGAGTAC GGCGACAAGT	160
	A V F V G S	A Q G I A K V	P P G P N I T	A E Y G D K	
		20		30	
161	GGCTGGACGC GAAGAGCACC	TGGTACGGCA AGCCGACGGG	CGCGGGGCCC AAGGACAACG	GCGGCGCGTG CGGGTACAAG	240
	W L D A K S T	W Y G K P T G	A G P K D N G G A C	G Y K	
		40		50	
241	GACGTGGACA AGGCCCCCTT	CAACGGCATG ACCGGCTGCG	GCAACACCCC CATCTTCAAG	GACGGCCGGG GCTGCGGCTC	320
	D V D K A P F	N G M T G C	G N T P I F K	D G R G C G S	
		60		70	
321	CTGCTTCGAG CTCAAGTGCT	CCAAGCCCGA GTCCTGCTCC	GGCGAGCCCA TCACCGTCCA	CATCACCACG GACAAAGAGG	400
	C F E L K C	S K P E S C S	G E P I T V H I T D	D N E	
		80		90	
401	AGCCCATCGC CCCCTACCAC	TTCGACCTCT CGGGCCACGC	ATTCGGGTCC ATGGCCAAGA	AGGGCGAGGA GGAGAACGTC	480
	E P I A P Y H	F D L S G H A	F G S M A K K G E E	E N V	
		100		110	
481	CGCGGCGCCG GCGAGCTGGA	GCTCCAGTTC CGCCGCGTCA	AGTGCAAGTA CCCGACGGC	ACCAAGCCGA CCTTCCACGT	560
	R G A G E L E	L Q F R R V	K C K Y P D G	T K P T F H V	
		120		130	
561	CGAGAAGGGA TCCAACCCCA	ACTACCTGGC GCTGCTCGTC	AAGTACGTCG ACGGCGACGG	CGACGTCGTG GCCGTGGACA	640
	E K G S N P	N Y L A L L V	K Y V D G D G	D V V A V D	
		140		150	
641	TCAAGGAGAA GGGCAAGGAC	AAGTGGATCG AGCTCAAGGA	GTCGTGGGGC GCCATCTGGA	GGATCGACAC CCCCAGACAAG	720
	I K E K G K D	K W I E L K E	S W G A I W R I D T	P D K	
		160		170	
721	CTCACCAGCC CATTCACCGT	CCGCTACACC ACCGAGGGCG	GCACCAAAGC AGAATTCGAG	GACGTCATCC CTGAGGGATG	800
	L T G P F T V	R Y T T E G	G T K A E F E	D V I P E G W	
		180		190	
801	GAAGGCCGAC ACCCATGACG	CATCCAAGTG AAGCCATCCT	GCCAGGAGAA GCCATCTTCT	TTCGGCCACC TTAATTAATT	880
	K A D T H D	A S K *			
		200		210	
881	TTCTCTAAGA TCATAAATAA	ACCAGCCGCA CAAATATATG	CATCCCAGTA CATGCAAATC	TCCTGTATGC ATGAATATAC	960
961	TCGTTTCATGC TGTAAGAGAG	GGAGAGATGC GAGGCTAATA	ATTTGAGGTC GAGACCATGT	CACCATGGGA GGTAAGTGA	1040
1041	GGAGTAATAC AAATTTTATA	TATCTCAAAA AAAAAAAAAA	AAAAAAAAA		1088

Fig. 2. cDNA and deduced amino acid sequences of Pha a 1 from canary grass pollen. The deduced amino acid sequence represented by the single letter code is shown below the DNA sequence, and begins at the first potential in-frame initiation codon indicated by numbers above the DNA sequence. The numbers on the sides refer to the numbering of the DNA sequence. The open reading frame continues for 269 amino acids and ends with a TGA stop codon denoted by an asterisk. The putative signal peptide is indicated by negative numbers. Underlined amino acid residues were identified by protein microsequencing [4].

26.2 kDa and a pI of 6.6 (according to MacProMass 2 Program). The amino acid composition from the deduced sequence is rich in the amino acid residues glycine (30 residues) and lysine (26 residues). There is a single potential *N*-glycosylation site having the characteristic Asn-X-Ser/Thr motif (amino acid residues 9–11) in the hydrophilic region of the mature protein. All these data suggest that clone 6 contains the full-length coding region.

Comparison of sequences with DNA databases

A search of existing DNA databases showed significant nucleotide identities between clone 6 and other group 1 grass pollen allergens. For example, there is a 78.9% identity between the nucleotide sequence of clone 6 and Lol p 1 from rye-grass pollen [6]. Similarly, there is a 77.1% identity between the nucleotide sequence of clone 6 and Hol l 1 from velvet grass pollen [26]. On the other hand, there is only 43.6% nucleotide sequence identity between clone 6 and Zea m 1 from maize pollen [27].

Comparison of sequences with protein databases

The amino acid sequence searches of existing protein databases indicated high amino acid identities between the deduced amino acid sequence of clone 6 and other group 1, 2 and 3 grass pollen allergens (Table 2). For example, there is 88.8% identity between the amino acid sequences of clone 6 and Lol p 1, 88.1% between clone 6 and Hol l 1 and 86.6% between clone 6 and Phl p 1 [28]. Moreover, the deduced C-terminal amino acid sequence of clone 6 shows 82.1% identity with that of the 28 amino

acid sequence identified as the major allergenic epitope of Lol p 1 [29].

Significant amino acid identities between clone 6 and other group 1 grass pollen allergens includes 63.8% with Cyn d 1 [20], 62.1% with Sor h 1 [30] and 52.8% with Zea m 1 [27]. Moreover, some degree of amino acid identity also exists between clone 6 and group 2/3 grass pollen allergens including 45.1% with Phl p 2 [31], 35.1% with Lol p 2 [32] and 28.9% with Lol p 3 [33]. It is interesting to note that the deduced amino acid sequence of clone 6 exactly matches that of *N*-terminal amino acid sequence determined for Pha a 1 previously [4]. All these comparisons suggest that clone 6 contains the full-length coding region of Pha a 1 from canary grass pollen.

Isolation and characterization of cDNA clones encoding polymorphic forms of Pha a 5

Five of the seven clones initially chosen for further study (e.g. clones 5, 14, 23, 28, 29) were subcloned into pBluescript plasmid vector for sequencing. Initial partial DNA sequencing revealed that these clones encode group 5 allergens based on sequence similarities to other group 5 allergens. A number of oligonucleotides, based on the homologous regions of the internal partial sequences of these clones, were synthesized and used as sequencing primers to obtain complete sequences of the cDNA clones (Table 1). The nucleotide sequences of clones 14 and 23 were identical to each other, however, clone 23 possessed an incomplete 3'-untranslated end and thus clone 14 was chosen for further sequence characterization. Nucleotide sequence of the other clones (e.g. clones 5, 28, 29) were found to be different both with each other and also with clone 14. For

Table 2. Percentage identity between the deduced amino acid sequence of clone 6 (Pha a 1) and other Group 1, 2 and 3 grass pollen allergens. References are represented by the numbers in square brackets

Grass name (Genus, species)	Allergen	% amino acid sequence identity with Pha a 1
Rye grass (<i>Lolium perenne</i>)	Lol p 1 [6]	88.8
Velvet grass (<i>Holcus lanatus</i>)	Hol l 1 [26]	88.1
Timothy grass (<i>Phleum pratense</i>)	Phl p 1 [28]	86.6
Bermuda grass (<i>Cynodon dactylon</i>)	Cyn d 1 [20]	63.8
Johnson grass (<i>Sorghum halepense</i>)	Sor h 1 [30]	62.1
Maize (<i>Zea mays</i>)	Zea m 1 [27]	52.8
Timothy grass (<i>Phleum pratense</i>)	Phl p 2 [31]	45.1
Rye grass (<i>Lolium perenne</i>)	Lol p 2 [32]	35.1
Rye grass (<i>Lolium perenne</i>)	Lol p 3 [33]	28.9

example, nucleotide sequence of clone 14 possessed 70.8%, 65.6% and 60.5% identity with that of clones 5, 28 and 29, respectively. Similarly, clones 28 and 29 shared significant nucleotide sequence identity (80.8%). However, the amino acid sequences of all clones differed significantly with each other (Fig. 3). The deduced amino acid sequence of clones 14, 28 and 29 indicated their highest amino acid identities with clone 5 (52%, 81.7% and 71.4%, respectively). Clone 29 was most similar to clone 28 with an amino acid sequence identity of 68.7%. However, clone 5 did not possess the 5' region and was not a full-length clone. Although not full-length, clone 5 possessed very high nucleotide and amino acid sequence identities with all other three clones, especially with clone 28 (95.5% nucleotide and 81.7% amino acid sequence

identity). The nucleotide sequences of these clones ranged in length from 760 to 1227 nucleotides.

Nucleotide and deduced amino acid sequence analysis

The longest open reading frame (ORF) of 960 nucleotides and 320 amino acid residues was encoded by clone 28. The ATG initiation codon of clone 28 is located at nucleotide position 28 and stop codon at position 987. The ORF corresponds to a polypeptide of 320 amino acid with an estimated molecular mass of 33.5 kDa (according to MacProMass 2 Program). The proposed translation initiation site and its flanking sequences (nucleotides 24–32) share 89% identity with the consensus plant sequence AACAAATGGC [24]. The most

Clone 28	MAVQKYTMALFLAVALVAGPAAP-----TPPTPRTPPLPPPRARDKATLTSRSE---	51
Clone 29	MAVQKYTVALFLAMALVAGPAASYAADAGTPPTPATPAV-PGAAAGKATTHEQKLIE---	56
Clone 14	MAVQKYTVALFLAVALVAGPAALYAGDGYAPATPAASATLATPATPAAS PQHAGTTEYHI	60
Clone 5	-----	
Common AA		
Clone 28	-----DINAASRRPWW-ASVPPADKFKTFADHVLCPNADV TSAATKAPQLKAKLDAA	103
Clone 29	-----DINAAFK--WWPASAPPADKYKTFETA---FSKANIAGASTKG-----LDAA	98
Clone 14	VRKAGLNEEKNAARQTDDEQKRSDEINCPDFNKS VHC--RADRLPVCSSSTSAHSSKQDVA	118
Clone 5	-----	
Common AA		
Clone 28	YRVAYEAAEGSTPEAKYDAFIAALTEALRVIAGAFEVHAVKPATEE-VVAD--PVGELQI	160
Clone 29	YSVVYNTAAGATPEAKYDSFVTALTEALRIMAGTLEVHAVKPATEEEVPSAKILRANSRS	158
Clone 14	WMLGYGSIQGFSD---DASVGSVSSEFHVIESAIEVITY---IGEEVKV--IPAGEVEV	170
Clone 5	-----AKYDAFIAALTEALRVIAGAFEVHAVKPATEE-VPAAKIPAGELQI	45
Common AA	* * * *	
Clone 28	VDKIDAAFKIAAT-----AANSAPANDKFTVFEGAFNKA I KESTAGAYETYKFIPS	211
Clone 29	STR-SSRFKIAATVATPLSHSTAANSAPANDKFTVFEGAFNKA I KERHGGPTETYKFIPS	217
Clone 14	INKVKAAFSTAA-----TAADEAPANDKFTVFVSSFNKA I KETTGGAYAGYKF IPT	221
Clone 5	VDKIDAAFKIAAT-----AANSAPANDKFTVFEGAFNKA I KERHGGAYETYKFIPS	96
Common AA	* * * * * * * * * * * * * *	
Clone 28	LEAA-VKQAYGATVARAPEVKYAVFEAGLT KAITAMSEA QKVAKP-PLSPQPPQVLPLAA	269
Clone 29	LEAA-VKQAYGATVARAPEVKYAVFEAGLT KAITAMSEA QKVAKPVRLSPQPPQVLPLAA	276
Clone 14	LEAA-VKQAYAASSATAPEVKYAVFETALKKA ISAMSEA QKEAKPA-----AAIS	270
Clone 5	LEASRSKQAYGATVARAPEVKYAVFEAGLT KAITAMSEA QKVAKP-----VRSVTAAAA	150
Common AA	*** * * * * * * * * * * * * *	
Clone 28	GGAATVAAASDVRVCRSHGTLQDACLLRCRGGCQPVVWRGGSHRARGGYKV	320
Clone 29	GGAATVAAASD-----SRGGYKV	294
Clone 14	AATTTISASTATPAAPPPPQLGTA-----TPAAVAGGYKV	305
Clone 5	GAATAAGGAATVAASRPTSA-----GGYKV	175
Common AA	*****	

Fig. 3. Comparison of deduced amino acid sequences of cDNA clones encoding group 5 allergens of canary grass pollen. Residues identical to all four proteins are indicated by an asterisk below the sequences. Dashes indicate gaps which have been inserted to give maximum alignment. Numbers to the right of the sequence refer to the numbering for each clone.

critical nucleotide, a purine at position -3 is conserved. The cDNA has a complete 3'-untranslated region that contains the canonical AATAAA polyadenylation signal [25] followed by a poly (A)⁺ tail.

The predicted protein sequence of clone 28 has a hydrophobic putative signal peptide sequence of 23 amino acids. This is indicative of a mature processed protein of 297 amino acids with an estimated molecular mass of 31.2 kDa and a pI of 9.7. The amino acid composition from the deduced sequence is rich in the amino acid residue alanine (67 residues). All these data suggest that clone 28 contains the full-length coding region.

Clone 29 has an ORF of 807 nucleotides starting with an initiation codon at nucleotide position 46 and terminating with a stop codon at nucleotide position 927. The ORF corresponds to a polypeptide of 294 amino acids with an estimated molecular mass of 30.3 kDa. The predicted protein sequence has a hydrophobic putative signal peptide sequence of 25 amino acids. This indicates a mature processed protein of 269 amino acids with an estimated molecular mass of 27.8 kDa and a pI of 10.3. The deduced amino acid sequence of this clone, as with other clones described above, is rich in amino acid residue alanine (66 residues).

Similarly, clone 14 has an ORF of 915 nucleotides starting with an initiation codon at nucleotide position 57 and terminating with a stop codon at nucleotide position 972. This ORF corresponds to a polypeptide of 305 amino acids with an estimated molecular mass of 31.4 kDa. The predicted protein sequence has a hydrophobic putative signal peptide sequence of 25 amino acids. This is indicative of a mature processed protein of 280 amino acids with an estimated molecular mass of 28.8 kDa and a pI of 6.3. The deduced amino acid sequence of this clone is rich in amino acid residue alanine (64 residues).

On the other hand, clone 5 is not a full-length clone as it does not have the initiation codon and lacks the leader and most of the other 5' sequence that other clones have. However, it does contain a terminating codon at nucleotide position 525.

Comparison of sequences with DNA databases

A search of existing DNA databases indicated significant nucleotide identities between clones 5, 14, 28 and 29 with other group 5 grass pollen allergens (Table 3). Nucleotide sequence of clone 28 possessed between 55 and 73.9% identity with other group 5 allergens (Table 3); the highest identity of 73.9% occurred with Lol p 5B from rye-grass pollen. Clone 14 possessed 72.7% nucleotide sequence identity with Poa p 5 clone 41 from KBG pollen. On the other hand, clone 29 indicated a 77.7% identity with nucleotide sequence of Phl p 5B from timothy grass. Although not a full-length clone, clone 5 displayed significant identities with nucleotide sequences of group 5 allergens of other grass pollens with the highest identity of 82.6% recorded with Phl p 5B (Table 3).

Comparison of sequences with protein databases

The amino acid sequence searches of existing protein databases indicated amino acid identities with other group 5 grass pollen allergens. Although lower than the nucleotide sequence identities, the level of amino acid sequence identities of clones 5, 14, 28 and 29 occurred, once again, with other group 5 allergens (Table 3). For example, the deduced amino acid sequence of clone 28 indicated 63.8% identity with the amino acid sequence of Lol p 5A. Clone 14's highest amino acid sequence identity of 57.4% occurred with Lol p 5B. On the other hand, the deduced amino acid sequences of both clones 29 and 5

Table 3. Percentage identity between the nucleotide and deduced amino acid sequences of clones encoding group 5 allergens of canary grass pollen and other group 5 grass pollen allergens. Open values refer to nucleotide sequence comparisons while values in parentheses refer to the deduced amino acid sequence comparisons. References are represented by the numbers in square brackets

Clone	% nucleotide (deduced amino acid) sequence identity						
	Lol p 5A [15]	Lol p 5B [16]	Phl p 5A [38]	Phl p 5B [38]	Poa p 5 (31) [12]	Poa p 5 (41) [12]	Poa p 5 (60) [12]
28	71.7 (63.8)	73.9 (51.9)	72.6 (59.3)	73.2 (57.9)	55 (49.1)	72.9 (54.7)	70.3 (55.4)
29	73.1 (65.9)	65.6 (57.8)	72.2 (54.4)	77.7 (58.9)	52.6 (52.7)	67.6 (60.2)	70.5 (56.5)
14	61.8 (49.8)	65.5 (57.4)	69.4 (48.4)	49.1 (45.7)	68.1 (45.9)	72.7 (53.1)	69.8 (51.5)
5	74.2 (80.0)	74.2 (62.3)	60.3 (66.9)	82.6 (74.3)	66.3 (68.6)	73.9 (70.9)	67.4 (69.1)

possessed their highest identities of 65.9% and 80%, respectively, with the amino acid sequence of Lol p 5A.

Discussion

Thirty-one clones have been isolated from a canary grass pollen cDNA library and out of the seven selected for further study, on the basis of strong human IgE antibody binding, two have been characterized as encoding Pha a 1 and five encoding Pha a 5. One of the cDNA clones (e.g. clone 6) encoded a full-length polypeptide of Pha a 1 while the other (e.g. clone 8) did not. With respect for cDNA clones encoding Pha a 5, two have been found identical (e.g. clones 14 and 23). However, only one (clone 14) encoded a full-length polypeptide. Similarly, both clones 28 and 29 encoded full-length polypeptides and clone 5 was only a partial clone. All of these cDNA clones were different, both at nucleotide and deduced amino acid level, with each other which encoded polymorphic forms of the canary grass pollen allergen Pha a 5.

Nucleotide and deduced amino acid sequence of Pha a 1

Full-length Pha a 1 was encoded by the cDNA clone 6. The ORF was 807 nucleotides and potentially encoded a protein of 269 amino acids with a 29 amino acids hydrophobic leader sequence. The molecular mass of the mature form of Pha a 1 (240 amino acids) was estimated to be 26.2 kDa. Thus, 22.9% of the 34 kDa molecular mass of natural Pha a 1 is probably due to post-translational modifications.

The cleavage site of the hydrophobic signal sequence can be identified by the start of the NH₂-terminal protein sequence which has previously been determined [4]. The -1 and -3 rule for the prediction of leader sequence cleavage site [34] which states that these positions should be small and neutral, holds for Pha a 1 with a -1: G and -3: A. Moreover, clone 6 also have the polyadenylation signal of AATAAA in the 3'-untranslated region, followed by a poly (A)⁺ tail [25].

The fusion-protein encoded by clone 6 has the capacity to bind IgE very strongly when tested with a serum pool of grass-pollen-allergic patients. This suggests that clone 6 encodes the major allergenic epitopes of Pha a 1. Moreover, the major IgE-binding epitope of Lol p 1 (YTTEGGTKSEFEDVIPEGWKADTSYSAK) to be located at the last 28 amino acids of the C-terminal region, shares significant identity (82.1%, with Pha a 1 (YTTEGGTKAEFEDVIPEGWKADTHDASK) [29]. This amino acid sequence is highly conserved among group 1 allergens of different grass pollens. Perhaps, this region of group 1 allergens are exposed continuously in the

context of the natural proteins for IgE binding. Furthermore, the existence of this highly conserved IgE epitope may explain the allergenic crossreactivity between closely related grasses, especially between Pha a 1 and Lol p 1 [4]. However, Cyn d 1 clones which expressed this region did not bind IgE from the sera of individuals allergic to Bermuda grass pollen [20]. A 15-mer peptide corresponding to this region did not bind IgE also. Comparison of the deduced amino acid sequence of the region in Cyn d 1 corresponding to the peptides isolated by Esch and Klapper [29] revealed 12 amino acid substitutions within the 28 amino acid sequence. This may explain the lack of allergenic crossreactivity among Cyn d 1 and other grass pollen allergens.

Homology between Pha a 1 and group 1 allergens of different grasses

The NH₂-terminal sequences for a number of group 1 allergens have been determined [5,20,35] and they show strong homology with that of Pha a 1 [4]. It has been suggested that group 1 allergens from all the temperate grasses contain hydroxyproline residues [7]. However, although NH₂-terminal sequences of Poa p 1, Phl p 1 and Sec c 1A contain hydroxyproline residues, none of the NH₂-terminal sequences of Pha a 1, Lol p 1 and Cyn d 1 do so [4].

The primary nucleotide and amino acid sequences of five group 1 allergens (Lol p 1, Cyn d 1, Hol l 1, Zea m 1, Phl p 1 and Sor h 1) have been previously determined [6,20,26-28,30, respectively]. A significant nucleotide identity was observed between clone 6 (Pha a 1) and Lol p 1 (78.9%) and Hol l 1 (77.1%). Comparison of the deduced amino acid sequence of Pha a 1 with other group 1 allergens revealed 88.8% identity with Lol p 1, 88.1% with Hol l 1, 86.6% with Phl p 1, 63.8% with Cyn d 1, 62.1% with Sor h 1 and 52.8% with Zea m 1. Moreover, we have also found that the entire 97 amino acid sequences of Lol p 2 [32] and Lol p 3 [33] have strong homology with the C-terminal end of Pha a 1. Pha a 1 has 35.1% identity with Lol p 2 and 28.98% with Lol p 3, while Lol p 2 and Lol p 3 have 59% identity with each other. A similar magnitude of amino acid identity has been previously reported between Lol p 1, Lol p 2 and Lol p 3 [6]. Moreover, the homologous C-termini of Lol p 1, Lol p 2 and Lol p 3 have been shown to contain crossreactive B-cell epitopes, binding to both human IgE and IgG antibodies [36]. It is interesting to note that the amino acid sequence of Phl p 2 possess 45.1% identity with that of Pha a 1. Now that the amino acid sequences of Pha a 1 and other group 1, 2, 3 grass pollen allergens are available, it should be possible to experimentally determine the T- and B-cell epitopes of

these allergens using synthetic peptides based on their sequences. Thus, presence of any crossreactive allergenic epitopes can be more precisely determined.

Nucleotide and deduced amino acid sequences of polymorphic forms of Pha a 5

The nucleotide and deduced amino acid sequences of three full-length and one partial Pha a 5 isoallergens have been determined in this study. None of the fusion-proteins encoded by the 31 clones isolated from canary grass pollen cDNA library bound the Lol p 5A-specific MoAb A7 suggesting that these clones encode different structures to that recognized by MoAb A7 of Lol p 5A. Four of the cDNA clones (e.g. clones 5, 14, 28 and 29) isolated have significant differences both at the level of nucleotide and amino acid sequences, suggesting that there are a number of genes encoding for isoforms of Pha a 5 belonging to the same multigene family. For example, two isoforms of the rye-grass pollen allergen Lol p 5 have been cloned and sequenced [15–16]. Crossreactivity studies using affinity purified IgE antibodies showed that both Lol p 5A and Lol p 5B share similar allergenic epitopes [16]. The two isoforms of Lol p 5 possess 66.4% amino acid sequence identities and are thought to be part of the same multigene family. Moreover, Lol p 1 has also been shown to exist in multiple isoforms [6,14]. These variants exhibit slight physicochemical differences but possess immunologically closely related properties. Thus, many major grass pollen allergens appear to exist as multiple isoforms and comparisons between the deduced amino acid sequences of these isoforms may help in elucidating the regions of the molecules that contain allergenic determinants.

Griffith *et al.* [10] identified four families of Amb a 1 allergens from short ragweed pollen. One particular plant was found to express seven unique Amb a 1 mRNA transcripts. Moreover, when the distribution of Amb a 1 isoforms in plants from different geographical locations (in the USA) were examined, there was no pattern to the distribution of isoforms. The canary grass pollen used to extract RNA for the cDNA library was collected from different plants occurring at different locations around the outer suburbs of Melbourne. The isoforms identified in this study might be a consequence of grass subspecies variation during pollen collection. However, whether these isoforms are expressed in the same plant or in different plants is not yet known.

Moreover, five cDNA clones encoding polymorphic forms of the major house dust mite allergen Der p 1 has been recently reported [17]. The translated sequences of all the five cDNA clones were found to be unique with

one to three differences between each clone and the polymorphisms included residues already described to be in T-cell epitopes. Thus, as well as being necessary to consider them in the construction of synthetic allergens, these substitutions will be important for the interpretation of experimental and genetic studies of pollen allergy.

Two-dimensional sodium, dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of crude grass pollen extracts reveals that major allergens exist as multiple isoforms occupying similar molecular mass but different pI. For example, out of the 17 IgE-binding proteins shown by two-dimensional Western analysis 12 belonged to variants of either of the two major allergens Lol p 1 and Lol p 5 [37]. Although two-dimensional gel and Western analysis of canary grass pollen crude extract has not been attempted, it may be highly likely that variants of Pha a 5 at different pI may exist, which some of them may be represented by the cDNA clones isolated in this study. The existence of these variants may suggest that they are encoded by genes that are part of the same multigene family.

Homology between Pha a 5 isoallergens and other group 5 allergens

The nucleotide and amino acid sequence searches of existing databases showed Lol p 5 (previously referred to as Lol p 1b, [15]), Poa p 5 (previously referred to as Poa p 9 [12]) and Phl p 5 [38] isoallergens have significant identity with clones 5, 14, 28 and 29. There was a 74.2% identity between the nucleotide regions of clone 5 and Lol p 5A. The amino acid sequence comparisons showed 80% identity between clone 5 and Lol p 5A. Similarly, clones 28 and 29 possessed 71.7% and 73.1% nucleotide identity (respectively) with Lol p 5A. The amino acid sequence comparisons indicated 63.8% and 65.9% identities between clones 28 and 29 (respectively) and Lol p 5A. On the other hand, clone 14 displayed 72.7% nucleotide (and 53.1% amino acid) identity with Poa p 5 allergen KBG 41.

The nucleotide and amino acid sequence comparisons between the clones isolated and sequenced in this study, and those already available in the databases, suggest that clones 5, 14, 28 and 29 isolated from a cDNA library of canary grass pollen encode polymorphic forms of a Pha a 5 allergen. These isoallergens all bound human IgE very strongly but demonstrated nucleotide and amino acid sequence differences, suggesting that their allergenic epitopes may still be conserved. These differences may be due to transcription from a number of different genes or different alleles of the same gene. Moreover, it will be interesting to investigate these differences further by synthetic peptides, based on their amino acid

sequences, in order to experimentally determine their crossreactive T- and B-cell epitopes. Moreover, synthetic peptides based on the amino acid sequences of these and isoforms of other grass pollen allergens may help in elucidating the regions of the molecules that contain the allergenic determinants.

Taken together, it is interesting to note that initial screening of the cDNA library indicated 31 positive clones from which only four were found to encode a potential group 1 allergen when the fusion-proteins were screened with group 1-specific MoAbs. The high frequency of cDNA encoding Pha a 5 rather than Pha a 1 may be due to the fact that mature canary grass pollen, which was used to construct the cDNA library with, may have contained more group 5 mRNA than group 1. Furthermore, Pha a 1 possessed higher nucleotide and amino acid identities with other group 1 allergens than Pha a 5 isoallergens did with other group 5 allergens. This suggests that some allergens are more conserved, perhaps through evolution, among botanically related grass pollens than others. Now that the amino acid sequences of these allergens are available, it should be possible to experimentally determine their T- and B-cell epitopes using synthetic peptides based on their sequences. The identified epitopes could then provide important information in the design of synthetic reagents for the diagnosis and treatment of grass pollen allergy.

Acknowledgments

The authors wish to thank Dr S. Krishnaswamy for supply of sera, the Australian National Health and Medical Research Council, the Asthma Foundation of Victoria and the University of Melbourne for a post-graduate scholarship to C. Suphioglu.

References

- Young P. Asthma and Allergies, an Optimistic Future. NIH Publ. No. 80-388, Washington, DC, 1980: 14-26.
- Smith JM. Epidemiology and natural history of asthma, allergic rhinitis, and atopic dermatitis (eczema). In *Allergy: Principles and Practice*. 3rd edn. E. Middleton (ed.) St Louis: C. V. Mosby Co., 1988:891-929.
- Smart IJ, Tuddenham WG, Knox RB. Aerobiology of grass pollen in the city atmosphere of Melbourne: effect of weather parameters and pollen sources. *Aust J Bot* 1979; 27:333-42.
- Suphioglu C, Singh MB, Simpson RJ, Ward L, Knox RB. Identification of canary grass (*Phalaris aquatica*) pollen allergens by immunoblotting: IgE and IgG antibody-binding studies. *Allergy* 1993; 48:273-81.
- Cottam GP, Moran DM, Standring R. Physicochemical and immunochemical characterization of allergenic proteins from rye-grass (*Lolium perenne*) pollen prepared by a rapid and efficient purification method. *Biochem J* 1986; 234:305-10.
- Griffith IJ, Smith PM, Pollock J et al. Cloning and sequencing of Lol p I, the major allergenic protein of ryegrass pollen. *FEBS Lett* 1991; 279:210-5.
- Matthiesen F, Nielsen AK, Søgaard TJ, Klysner S, Löwenstein H. NH₂-terminal sequences of four immunoaffinity purified grass pollen allergens: Phl p I, Poa p I, Sec c I and Cyn d I [abstract]. XVth European Congress of Allergology and Clinical Immunology, Paris, 10-15 May, 1992.
- Mole LE, Goodfriend L, Lapkoff CB, Kehoe JM, Capra JD. The amino acid sequence of allergen Ra5. *Biochem* 1975; 14:1216-20.
- Roebber M, Hussain R, Klapper D, Marsh D. Isolation and properties of a short ragweed pollen allergen, Ra6. *J Immunol* 1983; 131:706-11.
- Griffith IJ, Pollock J, Klapper DG, Rogers BL, Nault AK. Sequence polymorphism of Amb a I and Amb a II, the major allergens in *Ambrosia artemisiifolia* (short ragweed). *Int Arch Allergy Appl Immunol* 1991; 96:296-304.
- Nayak BN, Kisil FT. NH₂-terminal amino acid sequences of Poa p Ia and Poa p Ib allergens of Kentucky bluegrass pollen [abstract]. *J Allergy Clin Immunol* 1991; 87:325.
- Silvanovich A, Astwood J, Zhang L et al. Nucleotide sequence analysis of three cDNAs coding for Poa p IX isoallergens of Kentucky bluegrass pollen. *J Biol Chem* 1991; 266:1204-10.
- Johnson P, Marsh DG. Allergens from common ryegrass pollen (*Lolium perenne*) — I. *Immunochem* 1966; 3:91-100.
- Perez M, Ishioka GY, Walker LE, Chestnut RW. cDNA cloning and immunological characterization of the ryegrass allergen Lol p I. *J Biol Chem* 1990; 265:16210-5.
- Singh MB, Hough T, Theerakulpisut P et al. Isolation of cDNA encoding a newly identified major allergenic protein of rye-grass pollen: Intracellular targeting to the amyloplast. *Proc Natl Acad Sci USA* 1991; 88:1384-8.
- Ong EK, Griffith IJ, Singh MB, Knox RB. 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.
- 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.
- Staff IA, Taylor PE, Smith P, Singh MB, Knox RB. Cellular localization of water soluble, allergenic proteins in rye-grass (*Lolium perenne*) pollen using monoclonal and specific IgE antibodies with immunogold probes. *Histochem J* 1990; 22:276-90.
- Smart IJ, Heddle RJ, Zola H, Bradley J. Development of monoclonal mouse antibodies specific for allergenic components of ryegrass (*Lolium perenne*) pollen. *Int Arch Allergy Appl Immunol* 1983; 72:243-8.
- Smith P. Molecular analysis of Bermuda and rye-grass

- pollen allergens. PhD Thesis. 1993; The University of Melbourne, Australia.
- 21 Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987; 162:156–9.
 - 22 Messe E, Olson S, Leis L, Trent J. A quick method for high yields of lambda bacteriophage DNA. *Nucl Acid Res* 1990; 18:1923.
 - 23 Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 1977; 74:5463–7.
 - 24 Lutcke HA, Chow KC, Mickel FS et al. Selection of AUG initiation codon differs in plants and animals. *Embo J* 1987; 6:43–8.
 - 25 Birnstiel ML, Busslinger M, Strub K. Transcription termination and 3' processing: The end is in site! *Cell* 1985; 41:349–59.
 - 26 Schramm G, Bufe A, Becker W, Schlaak M. Characterization of major allergens of velvet grass pollen (Hol I I and Hol I V) by cDNA cloning [abstract]. *AN-EAACI Abs* 2215, *Allergy* 1993; 48:127.
 - 27 Broadwater AH, Bedinger PA, Rubinstein AL, Chay CH, Klapper DG. Zea m I, the maize homolog of the allergen-encoding Lol p I gene rye-grass. *Gene* 1993; 131:227–30.
 - 28 Laffer S, Valenta R, Vrtala S et al. Complementary DNA cloning of the major allergen Phl p I from timothy grass (*Phleum pratense*); recombinant Phl p I inhibits IgE binding to group I allergens from eight different grass species. *J Allergy Clin Immunol* 1994; 94:689–98.
 - 29 Esch RE, Klapper DG. Identification and localization of allergenic determinants on grass group I antigens using monoclonal antibodies. *J Immunol* 1989; 142:179–84.
 - 30 Avjioglu A, Creaney J, Smith PM et al. Cloning and characterization of the major allergen of *Sorghum halepense*, a subtropical grass. In: Kraft D, Schon A, eds. *Molecular biology and immunology of allergens*. Boca Raton: CRC Press 1993:161–4.
 - 31 Dolecek C, Vrtala S, Laffer S, Steinberger P, Kraft D, Scheiner O, Valenta R. Molecular characterization of Phl p II, a major timothy grass (*Phleum pratense*) pollen allergen. *FEBS Lett* 1994; 335:299–304.
 - 32 Ansari AA, Shenbagamurthi P, Marsh DG. Complete amino acid sequence of a *Lolium perenne* (perennial ryegrass) pollen allergen, Lol p II. *J Biol Chem* 1989; 264:11181–85.
 - 33 Ansari AA, Shenbagamurthi P, Marsh DG. Complete primary sequence of a *Lolium perenne* (perennial ryegrass) pollen allergen, Lol p III: comparison with known Lol p I and II sequences. *Biochem* 1989; 28:8665–70.
 - 34 von Heijne G. How signal sequences maintain cleavage specificity. *Mol Biol* 1984; 173:243–51.
 - 35 Matthiesen F, Schumacher MJ, Löwenstein H. Characterization of the major allergen of *Cynodon dactylon* (Bermuda grass) pollen, Cyn d I. *J Allergy Clin Immunol* 1991; 88:763–74.
 - 36 van Ree R, van Leeuwen WA, van den Berg M, Weller HH, Aalberse RC. IgE and IgG cross-reactivity among Lol p I and Lol p II/III. *Allergy* 1994; 49:254–61.
 - 37 Smith PM, Ong EK, Avjioglu A, Singh MB, Knox RB. Analysis of rye-grass pollen allergens using two dimensional electrophoresis and immunoblotting. In: Kraft D, Schon A, eds. *Molecular biology and immunology of allergens*. Boca Raton: CRC Press, 1993:141–3.
 - 38 Bufe A, Becker WM, Schramm G, Petersen A, Mamat U, Schlaak M. Major allergen Phl p V a (timothy grass) bears at least two different IgE-reactive epitopes. *J Allergy Clin Immunol* 1994; 94:173–81.

Copyright of Clinical & Experimental Allergy is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.