

Nucleotide sequence of cDNA encoding the group II allergen of cocksfoot/orchard grass (*Dactylis glomerata*), *Dac g II*

Roberts AM, Bevan LJ, Flora PS, Jepson I, Walker MR. Nucleotide sequence of cDNA encoding the group II allergen of cocksfoot/orchard grass (*Dactylis glomerata*), *Dac g II*.

Allergy 1993; 48: 615–623. © Munksgaard 1993.

Cocksfoot/orchard grass (*Dactylis glomerata*) anther cDNA clones encoding the group II allergen *Dac g II* were previously isolated on the basis of immunoreactivity of human, rabbit, and murine antibodies with a 24-kDa protein expressed as a fusion protein with β -galactosidase. Nucleotide sequencing reveals an open reading frame predicting expression of a 98-amino-acid (11-kDa) polypeptide exhibiting >90% homology with the group II allergen of *Lolium perenne*, *Lol p II*. *In vitro* translation of different sized clone fragments generated by polymerase chain amplification confirms eukaryotic expression of a 10–12-kDa polypeptide by SDS-PAGE and the position of a translational stop apparently unrecognized during expression of λ gt11 in *E. coli*. The unusual characteristics of the prokaryote-expressed fusion proteins may be exerting conformational alterations in *Dac g II*, as reflected by previous demonstrations of differences in human IgE immunoreactivity. Northern blot analysis using PCR-generated partial and full-length probes suggests that group II allergens may be encoded by a different family or families of temporally expressed genes from those encoding group I major allergens, although a group I gene may have been the progenitor.

A. M. Roberts¹, L. J. Bevan¹,
P. S. Flora¹, I. Jepson², M. R. Walker¹

¹ Molecular Biology Research Group, University Department of Clinical Biochemistry, Wolfson Research Laboratories, Queen Elizabeth Medical Centre, Birmingham, ² ICI Seeds (part of Imperial Chemical Industries plc), Jealott's Hill Research Station, Bracknell, Berkshire, UK

Key words: cDNA clones; cocksfoot/orchard grass; *Dactylis glomerata*; *Dac g II*; grass pollen allergens; nucleotide sequence; polymerase chain reaction.

Dr M. R. Walker
Molecular Biology Research Group
University Department of Clinical Biochemistry
Wolfson Research Laboratories
Queen Elizabeth Medical Centre
Birmingham B15 2TH
UK

Accepted for publication 22 February 1993

Immune recognition of grass pollen allergens contributes significantly to the conditions of seasonal allergic rhinitis and asthma (32). In temperate latitudes, pollens from cocksfoot/orchard grass (*Dactylis glomerata*: *Dac g*), rye grass (*Lolium perenne*: *Lol p*), timothy grass (*Phleum pratense*: *Phl p*), and Kentucky bluegrass (*Poa pratense*: *Poa p*) represent major allergen contributors. Understanding the molecular basis of B-cell and T-cell recognition of allergen structures, uniquely expressed and coexpressed within and between grass species, is fundamental to the development of many forms of immunomodulatory therapy (26). This depends, in part, upon the determination of inter- and intraspecies sequence similarities.

Immunochemical characterization of grass pollen allergens, notably from *Lol p*, *Dac g*, *Phl p*, and *Poa p*, has revealed the existence of several protein groups (16) bearing protein sequence and B-cell epitope similarities. Application of protein microsequencing and recombinant DNA technologies has contributed to modification of these classifications and has re-

vealed the existence of several isoforms within each allergen group (11, 15, 20, 22, 30, 31). Nevertheless, group I and V/IX glycoproteins (27–33 kDa) are regarded as major allergens, each present as several isoform products from each gene family, to which 85–90% of grass-allergic patients are sensitive (8). Group II and III allergens have also been characterized as immunologically distinguishable isoallergenic proteins (11 kDa) to which 40–45% of grass-allergic patients are sensitive (1). The latter groups of allergens have been regarded as having minor clinical significance despite, in the case of group II, their sequence homology with the C-terminal half of group I allergens. Similarities in immunologic responsiveness related to HLA haplotype (9) between *Lol p I*, II and III further suggest that group II and III allergens may be derived from breakdown of group I isoallergen gene products or that they are encoded by separate gene families which may have arisen from deletional and point mutation of an original group I gene.

We have previously described the identification of

cDNA clones in a λ gt11 library constructed from *D. glomerata* (*Dac g*) anther mRNA which lead to expression of an immunoreactive 24-kDa polypeptide fused to β -galactosidase (33), consistent with membership of group I or V/IX (17–19, 21, 24). Immunologic studies using human polyclonal, rabbit polyclonal, and murine monoclonal antibodies revealed that these clones encode *Dac g* II allergens. Furthermore, the recombinant fusion protein shared immunoreactivity with anti-*Lol p* I and II antibodies, and these data suggested an increase in human IgE reactivity with *Dac g* II within sera from patients receiving grass pollen immunotherapy (25). Here we report the complete nucleotide sequence of recombinant *Dac g* II clones exhibiting an open reading frame for a 98-amino-acid polypeptide (11 kDa aglycosylated). The extremely high degree of homology with the *Lol p* II amino acid sequence and Northern blot studies support the suggestion that group II allergens are encoded by a gene family distinct from group I allergens. The unusual nature of the inserts has also been investigated (i.e., expression of a 24-kDa polypeptide in λ gt11), and it provides a possible rationale for the apparent increase in IgE reactivity in some grass immunotherapy patient sera (25).

Material and methods

cDNA clones

Isolation of clones 7.8, 8.1, and 8.3, together with preparation of bacterial (*E. coli* strain Y1090) lysates has been previously described (25, 33). These clones all generate a 140-kDa fusion protein with β -galactosidase representing 24 kDa of insert encoded polypeptide recognized by human, rabbit, and rodent antibodies (25).

Oligonucleotides

Oligonucleotides were synthesized by standard phosphoramidate chemistries on a Biotech Instruments BT8500 DNA synthesizer by Dr J. E. Fox (Alta-Bioscience, Biochemistry Department, University of Birmingham, UK). After synthesis, oligonucleotides were diluted to a final concentration of 0.2 μ mol/ml in sterile, deionized water (SDW) and stored at -20°C . These were diluted 1 in 20 in SDW before use in amplification by polymerase chain reaction (PCR). Sequences of oligonucleotide primers used for *in vitro* amplification and nucleotide sequencing are given in Table 1, and hybridization sites are indicated in Figs. 1 and 2a.

Polymerase chain reaction (PCR)

Allergen clones (5- μ l aliquots) were added to 95 μ l SDW and boiled for 30 min. Aliquots (5 μ l) from the

Table 1. Oligonucleotide sequences used as primers for PCR amplification, sequencing, and RAWIT

Code	Sequence (5' $\rightarrow$ 3')	Application ¹
D222 ²	GGTGGCGACGACTCCTGGAGCCCG	A, S
D223 ²	TTGACACCAGACCACTGGTAATG	A, S
D2056	ATCAAGTACAACAAGGAG	S
D2805	CTCCTTGTTGTACTTGAT	S
D2819	CACGCATATAAGCACCTCAC	A, S, R
D2231	ACGGCGTGTGGGAGATCAA	S
D2806	TTGATCTCCACACGCCGT	A, S, R
D2351	GTCGGCACCACCTACAAG	A, S
D2057	GTCACCTATAGTGATACC	A, S, R
D2352	CTTGATAGTGGTGCCGAC	A, S, R
D2308 ³	TAATACGACTCACTATAGGGAGACCACCATGGAGG- CCCCGTGACCTTC	R

¹ Use in PCR probe/template amplification (A), nucleotide sequencing (S), and/or RAWIT (R).

² Specific for λ gt11 Lac z sequences flanking *EcoRI* cloning site.

³ RAWIT primer based on 5' addition of T7 polymerase promoter sequence and translational initiation sequence added to clone-specific sequence (underlined).

boiled clones were subsequently used as templates for *in vitro* amplification by PCR. Reactions were performed in a total reaction volume of 100 μ l, comprising 10 μ l 10 $\times$ reaction buffer (Cetus ILS, London, UK), 200 μ M (2.5 μ l) of each dNTP, 100 pmols (10 μ l) each of the appropriate 5' and 3' primers (Table 1 and Fig. 1), and 55 μ l SDW. After denaturation at 94°C for 7 min, 2 units (0.4 μ l) Ampli-taq[®] (Cetus ILS, London, UK) was added to each reaction, and samples were overlaid with light mineral oil. Samples that were to be used in sequencing reactions and probe generation were subjected to 25 temperature cycles of 94°C for 30 s, 45°C for 90 s, and 72°C for 120 s. Samples that were to be used in *in vitro* translation experiments (see below) were initially subjected to three temperature cycles of 94°C for 60 s, 40°C for 90 s, and 72°C for 180 s, followed by 25 temperature cycles of 94°C for 30 s, 50°C for 90 s, and 72°C for 180 s. Amplification products were assessed by electrophoretic separation of 10- μ l aliquots in 1% agarose gels and visualization of ethidium bromide stained bands under UV illumination.

For nucleotide sequencing and probe generation, the remaining 90 μ l of PCR reaction was purified by the MagicTM PCR Prep DNA purification system (Promega Corporation, Southampton, UK). Purified samples were freeze-dried in a Univapo 100 (Jones Chromatography, Cardiff, UK) and resuspended in 12 μ l SDW.

For *in vitro* translation experiments, amplified PCR products were isolated by MagicTM PCR Prep (Promega Corporation, Southampton, UK), as described above. However, the smallest PCR product (i.e. D2308 + D2806 – see Fig. 4a, track 5) was purified by electroelution onto Whatmann[®] 3MM

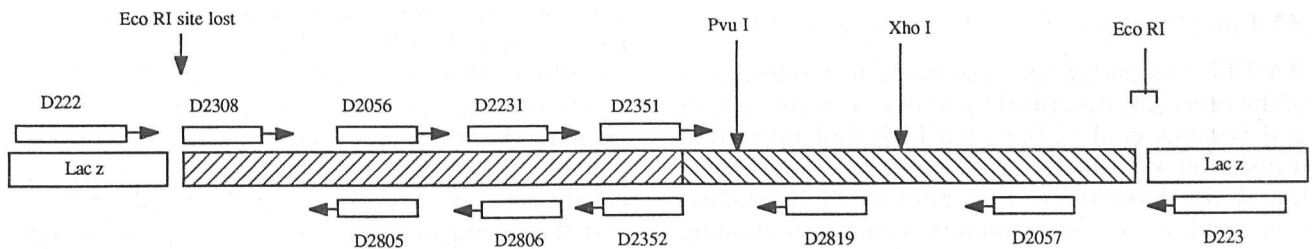


Fig. 1. Schematic representation of cDNA clone inserts showing partial restriction map and position of some oligonucleotide primer hybridization sites used for amplification and sequencing. Shading represents the different subfragments generated as probes for Northern blot hybridization.

chromatography paper, and concentrated by centrifugation (3000 g) in a Centricon 10™ microcon-

trator (Amicon, Basingstoke, UK) before freeze-drying.

15 10
GAGGCCCGCTGACCTTCACGGTGGAGAAGGGTTCCGACGAGAAGAACCTGGCG 541
E A P V T F T V E K G S D E K N L A

20 30
CTGTCTGATCAAGTACAACAAGGAGGGCGACTCCATGGCGGAGGTGGAGCTCAAG 108
L S I K Y N K E G D S M A E V E L K

40 50
GAGCAGCGCTCCAACGAGTGGCTCGCCCTGAAGAAGAACGGCGACCGCGTGTGG 162
E H G S N E W L A L K K N G D G V W

60 70
GAGATCAAGAGCGACAACCGCTCAAGGGCCCATTCAACTTCCGCTTCGTCTCC 216
E I K S D K P L K G P F N F R F V S

80 90
GAGAAGGGGATGAGGAACGTCTTCGTAGACGTGGTTCGGCTGACTTCAAGGTC 270
E K G M R N V F V D V V P A D F K V

98
GGCACCACCTACAAGCCCCGAGGAGTAGGGCGCCGCTTCCGCCAGCCGCCGG 321
G T T Y K P E E * A A A S A S R R
(E)

CGATCGTCGGAAGTTTTTCAATTCCTAATTTTATCTTGTCAAGGTCGTATCGTA 375
R S S E V F Q F L I L S C Q G R I V

AATAATTGTGAGGTGCTTATATGCGTGATGAGGAGAGGCAACGCGATGTGCTTG 429
N N C E V L I C V M R R G N A M C L

ATTGCTAGTATCTCCATGCACCACATCTTGACTCTTGATCGGTTCTTTTTTGAT 483
I A S I S M H H I L T L D R F F F D

GGACTCGAGATTATTTATAAAATATTCAAATGATGTTTCAAAAACCTCGTAGC 537
G L E I I Y K I F K M M F Q K P R T

AGAACTTGTATAGAGAAGGATTTCCCGCGCGCAGCAGTCTTCCATACCAACT 591
R T C I E K D F P R R S S S S I P T

TAGCTGCCGGCGCTGCTATTGGCATAACAACCGGTACACCATAGGTTGACCCA 645
*

ACCCAGTCTCTCGTACTAGGGTTGGCTCCTCGAGTTCTCCCCGAATTC 696

§ denotes amino acid numbering and ¶ denotes nucleotide numbering (of inserts)

Dac g II EAPVTFTEVEKGSDEKNLALS I KYNKEGDSMAEVELKEHGSNEWLAL
Lol p II A---E-----
A
Lol p III TK-DL-----A-T-V-N---TRP---TL-----RQ-----E---EPM
Lol p I^c DTKPT-H---NPNY--ILV--VDGDGDVVA-DI--K-KDK-IE-
A NPNY T
^c(starting from amino acid 144)

Dac g II KKNNGDGVWEIKSDKPLKGFNFNRFVSEKGMNRNVFVDVVPADFKVGT
Lol p II -----D-----
D
Lol p III T-K- NL-----A---T---M---L-KG-----DE-I-TA-T---K
Lol p I -ESWGA--R-DTPDK-T---TV-YTT-G-TKSEVE--I-EGW-AD-
F

Dac g II TYKPEE
Lol p II -----
Lol p III --T--YN
Lol p I S-SAK

C

Dac g II GAGGCCCGTGACCTTCACGGTGGAGAAGGGTCCGACGAGAAGAACCTG
Lol p I --CA--AAGCCA--G---CAC--C-----A---CCC--CT-----

Dac g II GCGCTGTCGATCAAGTACAACAAGGAGGCGACTCCATGCGCGAGGTGGAG
Lol p I --TA-TCT-G-G-----GT-G-C-GC-A--G-GA-G---T--CC-----C

Dac g II CTC AAGGAGCAGCGGTCCAACGAGTGGCTCGCCCTGAAGAAGACGGCGAC
Lol p I A-----A-G---AAGG-TA-----A---AG--C---G--TCGT-G-GA

Dac g II GGCGTGTGGGAGATCAAGAGCGACAAACCGCTCAAGGGCCCATTC AACTTC
Lol p I -CA--C---AG----G-C-C-CC-G-TAA---G-C-----C-G--

Dac g II CGCTTCGTCTCCGAGAAGGGGATGAGGAACGCTCTTCGTAGACGTGGTCCG
Lol p I ---A-AC-A-----GGC--C-CC-AATC--AAG--AG--T--CA-----C

Dac g II GCTGACTTCAAGGTGCGGACCACTACAAGCCCGAG GAGTA
Lol p I -AG-G--GG-----C--A---TT-----TC-G--A--T---C-

Fig. 2. a) *Dac g II* cDNA clone nucleotide and inferred amino acid sequences. Primer sequences are underlined (see also Table 1), and position of 3' *EcoRI* site is shown in boldface. Position of stop codon is indicated by asterisk and is followed by possible translation (in *italics*) of remainder of reading frame (see text). b) Amino acid sequence homologies between *Dac g II*, *Lol p II*, *Lol p III*, and *Lol p I*. Dashes represent identity, and amino acids in bold represent isoallergen differences. Gap has been introduced in *Lol p III* sequence for maximal alignment. c) Comparison of *Dac g II* nucleotide sequence with *Lol p I* (3' coding region).

RNA amplification with in vitro translation (RAWIT)

RAWIT was performed, according to modification of the protocols described by Sarkar & Sommer (28) and Sommer et al. (29), with a T7-linked eukaryotic translation kit (Promega Corporation, Southampton, UK). Essentially, 1 µg purified PCR product was used in a 25-µl reaction volume containing 12.5 µl TNT lysate, 1 µl buffer, 0.5 µl TNT RNA polymerase, 0.5 µl amino acid mix minus leucine, 0.5 µl RNasin, 6.5 µl SDW, and 2.5 µl ^{14}C -leucine (50 µCi/ml $^{-1}$) (Amersham International plc). Reactions were incubated at 30°C for 90 min, and labelled translation products were assessed by electrophoresis of 10-µl aliquots in 17% cross-linked SDS-PAGE (14). Gels were dried and subjected to autoradiography for up to 7 d at room temperature, with Hyperfilm β-max film (Amersham International plc).

Nucleotide sequencing

Direct nucleotide sequencing of PCR products was performed as previously described (23). Thus, 2 µl sequencing primer, 2.5 µl concentrated PCR fragment, 1 µl 6× reaction buffer (240 mM Tris/HCl pH 7.0, 150 mM MgCl₂, 300 mM NaCl, and 60% v/v dimethyl sulphoxide (DMSO)), and 0.5 µl single-stranded binding protein (SSB) (United States Biochemicals) were mixed and heated to 96°C for 3 min. Amounts of 1 µl 100 mM dithiothreitol and 1 µl ^{35}S -dATP-αS (10 µCi/µl $^{-1}$) (Amersham International plc) were then added, mixed, and immediately frozen in liquid nitrogen. Each reaction tube was removed in turn, and when the mix was almost thawed, 2 µl of Sequenase® V2.0 (United States Biochemicals) previously diluted 1 in 8 in SDW was added. After mixing, 2-µl aliquots were immediately added to 2 µl of each of the four termination mixes, previously made 10% for DMSO. These were incubated for 5 min at 37°C, after which 2 µl chase mix (containing 0.25 mM dNTPs, 150 mM NaCl, and 10% DMSO) was added, and reactions were incubated for a further 15 min before the reactions were stopped. Proteinase K was then added (0.2 µg/µl $^{-1}$) and incubated at 60°C for 20 min. Extension termination products were separated under denaturing gel electrophoresis in urea/polyacrylamide (27) and subjected to autoradiography overnight with Hyperfilm-β-max (Amersham International plc) at room temperature.

Sources and extraction of RNA

Total RNA was extracted from frozen *Dactylis* grass heads (collected at or after antherogenesis) and pollen (Greer Labs, NC, USA) by grinding the tissue in

a pestle and mortar, chilled with liquid nitrogen. Extraction buffer (50 mM Tris pH 9.0, 0.2 M NaCl, 1.0 mM EDTA, 1.0% SDS) was added, and grass heads were homogenized on ice for 4×10 s. The resulting homogenate was mixed with 5 ml acid phenol/chloroform/isoamyl alcohol (25:24:1), sonicated on ice for 4×20 s, mixed by hand for 5 min, and then centrifuged at 7650 g for 15 min at 4°C. The aqueous phase was removed and was re-extracted three more times. RNA was precipitated overnight with 0.5 ml 3 M NaOAc, 10.0 ml 100% EtOH, at -20°C. After centrifugation at 7650 g for 20 min at 4°C, RNA pellets were washed with 70% EtOH, resuspended in 2 ml of diethyl pyrocarbonate (DEPC)-treated water, and precipitated overnight with 2 ml of 4 M LiCl at approximately 1°C. Precipitated RNA was centrifuged at 7650 g for 30 min at 4°C, resuspended in 1 ml of DEPC-treated water, and precipitated with 1 µl 20 mg/µl glycogen, 0.5 vol 3 M NaOAc, and 2 vol EtOH on dry ice for about 2 h. RNA was pelleted by centrifugation at 8000 g for 30 min at 4°C, washed with 70% EtOH, and resuspended in 20 µl of DEPC-treated water.

Total RNA was also obtained from a maize inbred line by a previously described method (13).

Northern blotting

After electrophoretic separation on 1.2% formaldehyde agarose gels, RNAs (20 µg) were blotted onto Hybond N membrane (Amersham International plc) (27). PCR products generated with primers D222 and D223 (full insert), D222 and D2352 (5' half of insert), and D223 and D2351 (3' half of insert) (Fig. 1) were purified by Gene CleanTM (full insert) or the MERmaidTM purification procedure. Probes (50 ng) were radiolabelled with α- ^{32}P -dCTP (10 µCi/µl $^{-1}$) (Amersham International plc) by the random primer method (7) and used in Northern hybridization in 5× SSPE, 1% blotto, 0.5% SDS, and 200 µg/ml salmon testis DNA at 65°C for 18 h. After washes at 0.1× SSC 0.1% SDS, filters were exposed to Hyperfilm-MP (Amersham International plc) for 2–24 h. Control hybridizations were performed with either the pollen-specific probe Zmc13 (12) or a wheat ribosomal probe (10) to show that equal RNA loadings had been achieved.

Results*Nucleotide sequencing*

The nucleotide sequence of clones 7.8, 8.1, and 8.3 was obtained by dideoxy sequencing of PCR-amplified inserts in both directions with primers designed to give overlapping sequence data (Table 1 and Fig. 1). The complete nucleotide and inferred amino acid sequence of recombinant *Dac g II* are

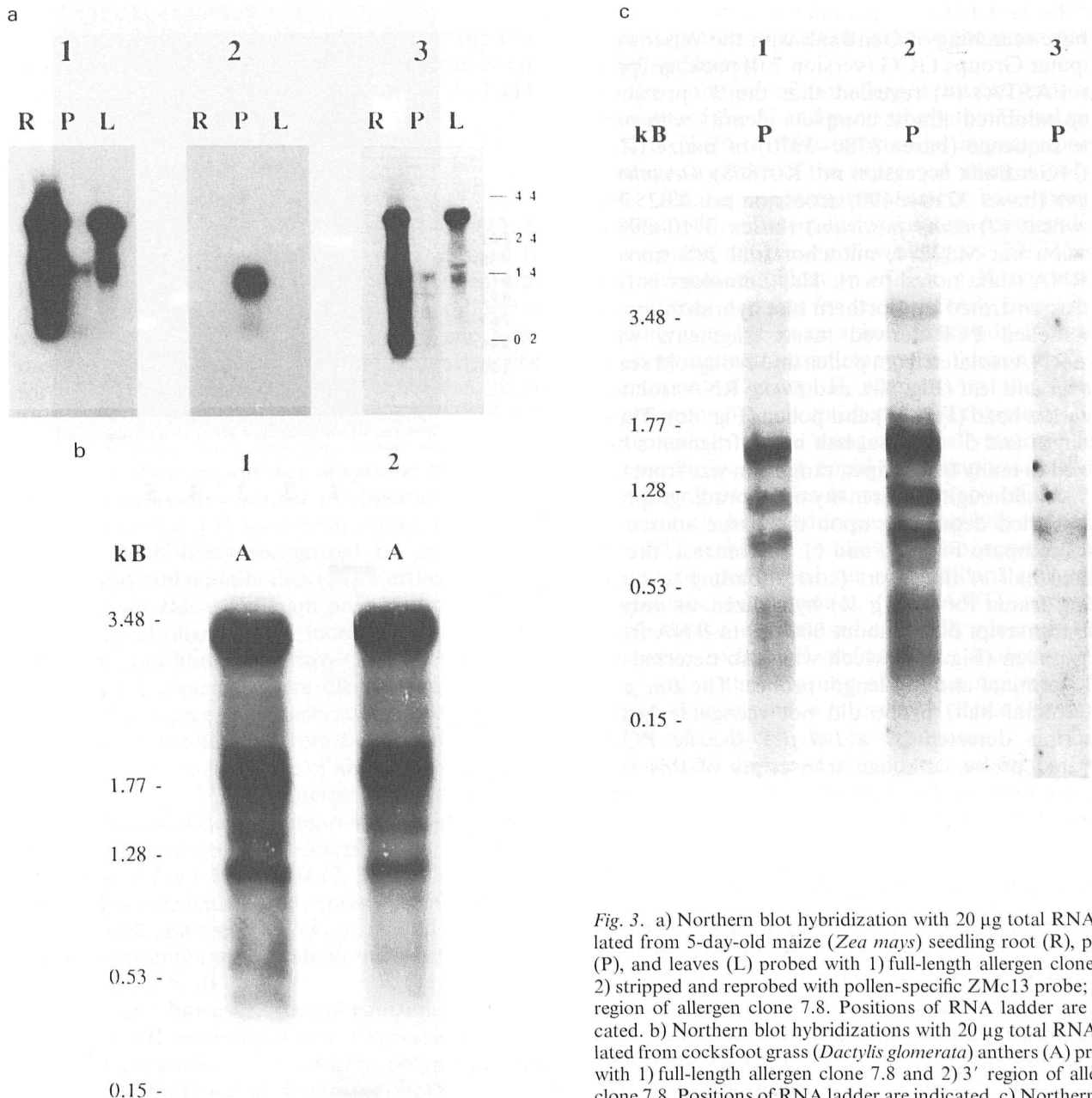


Fig. 3. a) Northern blot hybridization with 20 μ g total RNA isolated from 5-day-old maize (*Zea mays*) seedling root (R), pollen (P), and leaves (L) probed with 1) full-length allergen clone 7.8; 2) stripped and reprobed with pollen-specific ZMc13 probe; 3) 3' region of allergen clone 7.8. Positions of RNA ladder are indicated. b) Northern blot hybridizations with 20 μ g total RNA isolated from cocksfoot grass (*Dactylis glomerata*) anthers (A) probed with 1) full-length allergen clone 7.8 and 2) 3' region of allergen clone 7.8. Positions of RNA ladder are indicated. c) Northern blot hybridizations with 20 μ g total RNA isolated from cocksfoot grass (*Dactylis glomerata*) pollen (P) probed with 1) full-length allergen clone 7.8, 2) 3' terminal fragment, and 3) 5' terminal fragment. Positions of RNA ladder are indicated.

shown in Fig. 2a. The insert contains 693 bp containing restriction sites for *Pvu* I and *Pst* I flanked by a 3' *Eco*RI site but which lacked the 5' *Eco*RI site used in the original cloning procedure (33). This is consistent with the inability to remove the insert by restriction of the vector with *Eco*RI (unpublished results). The insert also contains an open reading frame encoding a 98-amino-acid polypeptide. Comparison of the inferred amino acid sequence of this 98mer with the amino acid sequences of *Lol p* I (from residue 144), *Lol p* II, and *Lol p* III (Fig. 2b) shows

that this portion of the insert exhibits maximum homology (>90%, accounting for isoallergen differences) with *Lol p* II. These amino acid differences may be accounted for by single (E→A, S→A, N→D) or double (T→E, V→D) base substitutions. The predicted amino acid sequence of this open reading frame portion of the inserts exhibits approximately 60% homology with *Lol p* III, and approximately 32% homology with the C-terminal portion of *Lol p* I (Fig. 2b), while there is approximately 60% homology with *Lol p* I at the nucleotide level (Fig. 2c).

Of the remaining 3' terminal half of the insert, database searching of GenBank with the Wisconsin Computer Groups GCG (version 7.0) package (program FASTA) (4) revealed that the 3' proximal 190 bp exhibited almost complete identity with part of the sequence (bases 3780–3970) of maize (*Zea mays*) (GenBank accession no. K01868), *Oenothera lutea* (bases 3210–3400: accession no. X02559), and wheat (*Triticum aestivum*) (bases 3710–3900: accession no. M37274) mitochondrial 26S ribosomal RNA (data not shown). This homology is further demonstrated by Northern blot hybridization of radiolabelled PCR-derived insert fragments with maize RNA isolated from pollen and 5-day-old seedling root and leaf (Fig. 3a), and *Dac g* RNA isolated from grass head (Fig. 3b) and pollen (Fig. 3c). Thus, full-length and 3' terminal half insert fragments hybridized to many transcripts, ranging in size from 0.2 to 3.5 kb, although the intensity of autoradiographed bands varied depending upon the tissue source of RNA (compare Figs. 3b and c). In contrast, the 5' terminal half of the insert (corresponding to open reading frame for *Dac g* II) hybridized to only a single transcript band (about 800 bp) in RNA from *Dac g* pollen (Fig. 3c), which was also detected by the 3' terminal and full-length probes. The *Dac g* II (5' terminal half) probe did not recognize larger transcripts detected by a *Lol p* I specific PCR-generated probe, although transcripts of this size were also detected by full-length and 3' terminal half probes (data not shown).

In vitro translation of insert subfragments

Subfragments of the cDNA clones were generated by PCR with a 5' terminal T7 polymerase adaptor modified forward primer (D2308) in combination with reverse primers hybridized to different portions of the insert (Figs. 1 and 4a). *In vitro* translation of these subfragments in a T7-linked eukaryotic translation system confirmed the location of the translational stop codon identified by nucleotide sequencing. Thus, PCR products generated by D2308 in combination with D2352, D2057, and D223 (hybridizing at or downstream of the stop codon) all generated a 10–12-kDa polypeptide, as determined by SDS-PAGE/autoradiography of incorporated

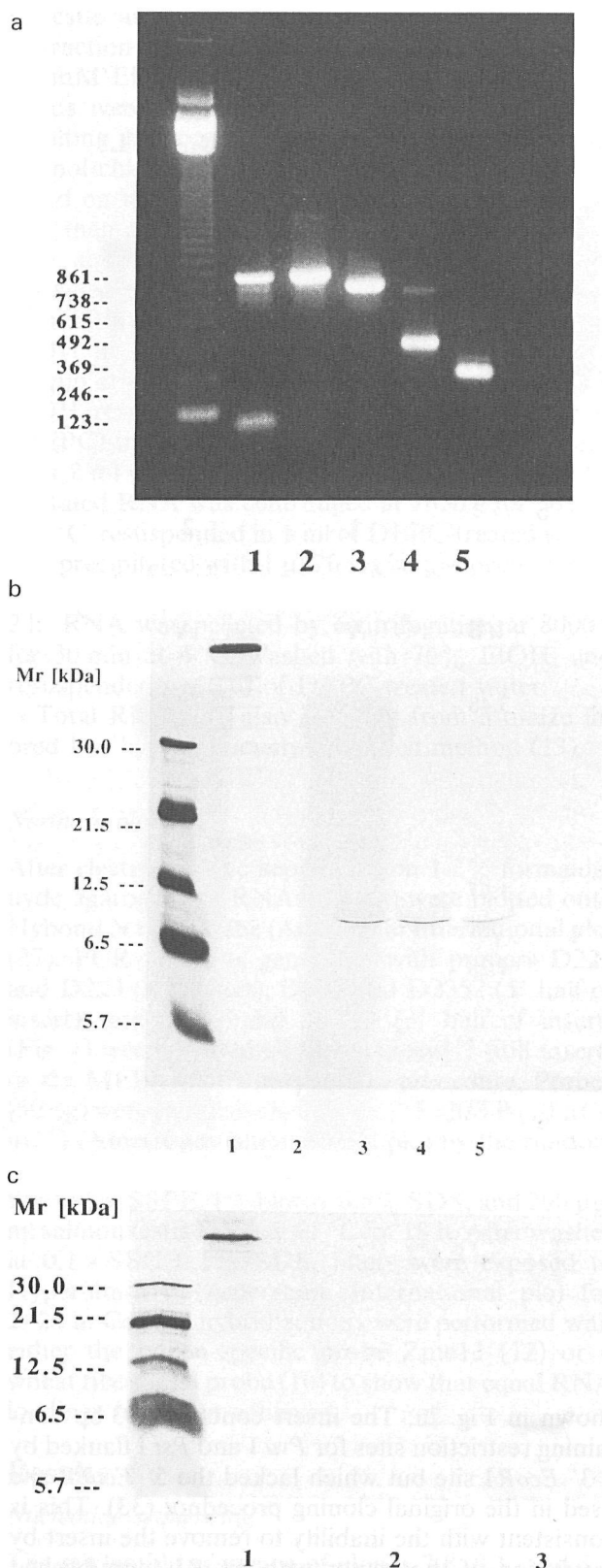


Fig. 4. a) 1% agarose separation of PCR products (10 μ l) generated from cDNA clone 8.1 with primers 1) D222 + D223, 2) D2308 + D223, 3) D2308 + D2057, 4) D2308 + D2352, and 5) D2308 + D2806 (see Fig. 1). Positions of 123 bp markers are indicated. b) Autoradiograph of ^{14} C-leucine-labelled *in vitro* translation products, separated by 17% cross-linked SDS-PAGE, using 1) positive control luciferase plasmid DNA, 2) negative control (no DNA template) and PCR products generated with primers, 3) D2308 + D223, 4) D2308 + D2057, and 5) D2308 + D2352.

Positions of mol. mass markers are indicated. c) Autoradiograph of ^{14} C-leucine-labelled *in vitro* translation products, separated by 17% cross-linked SDS-PAGE, using 1) positive control luciferase plasmid DNA, 2) negative control (no DNA template), and 3) PCR product generated with primers D2308 + D2806. Positions of mol. mass markers are indicated.

^{14}C -leucine (Fig. 4b, lanes 3–5). In contrast, PCR products corresponding to the 5' half of the ORF portion of the insert, generated by D2308 in combination with D2806, gave an expected 6–7-kDa-labelled polypeptide (Fig. 4c, track 3).

Discussion

Cloning of full-length transcripts of *Lol p I* (11, 22) and *Lol p V/IX* (31) allergens has confirmed the presence of isoallergens among grass pollen allergens, and sequence and cell targeting differences suggest that these represent separate gene family members. Similar isoallergenicity has been described for group V/IX proteins from other grass (15, 20, 30) and weed species. These data have a bearing on immunotherapy since they may reflect differences in the timing of the appearance of specific allergen gene products expressing similar or dissimilar epitopes. *Lol p I* and *Dac g II* have been shown to be cross-reactive, both being recognized by cross-reactive human and rabbit antibodies (25). Furthermore, with recombinant *Dac g II* fusion proteins, it is possible that group II allergens are increased in recognition after allergoid immunotherapy (25).

The data presented here clearly demonstrate a remarkable degree of conservation in sequence between group II allergens from *Dac g* and *Lol p*, and a consequently high degree of similarity with *Lol p I* and III (2, 3, 11, 22). It is interesting to note that *Dac g II* cDNA encodes 98-amino-acid residues rather than the 97 amino acids described by peptide sequencing of *Lol p II* and III (2, 3). The sequence of *Dac g II* is consistent with expression of similar putative T-cell epitopes on *Lol p I*, II, and III and the consequent immunologic similarities between these molecules (1, 5, 6).

There are, however, several unusual features in these *Dac g II* encoding clones. Expression in *E. coli* leads to generation of a 24-kDa polypeptide (33) despite the presence of a translational stop codon recognized by eukaryotic *in vitro* translation experiments (Figs. 4b and c). This cannot be explained by *E. coli* reading of the UAG codon as glutamine by SupE⁺ amber suppressor mutations within the host, since the lysates were prepared in the SupE[−] host strain Y1090 (27). It is possible that the stop codon is obscured by significant secondary structures in this portion of the sequence because of locally high concentrations of G and C bases. Indeed, sequencing reactions exhibited a high degree of compression in this region (data not shown). It is conceivable that this leads to expression of a polypeptide including the residues shown in italics in Fig. 2a. We are currently addressing these questions by the RAWIT approach to generate C-terminal fragments with T7-promoter modified primers, on the basis of hybrid-

ization to a site immediately downstream of the stop and staggered to read on each of the three possible reading frames.

The addition of nonallergen fusion partner sequences to the 5' and particularly the 3' end of *Dac g II* is likely to have substantial impact upon the tertiary structure of the expressed allergen polypeptide. Similar conformation effects have been described with different fusion protein partners linked to recombinant grass pollen allergens (22, 31). Conformational modification of *Dac g II* expressed in *E. coli* may account for its apparently increased reactivity with IgE in sera from grass pollen immunotherapy patients, mimicking the "denatured" conformation of group II allergens in the allergoid preparations. Such an altered reactivity may not be detectable in assays using essentially native allergen preparations, but it could have important implications for grass pollen immunotherapy efficacy. This is currently being addressed by optimization of the RAWIT technique to allow immunoreactivity studies to be performed on recombinant polypeptide fragments. Furthermore, the RAWIT system described here may be used in rapidly generating small quantities of *Dac g II* subfragments, analysis of which may aid in the targeting of certain regions for site-directed mutagenesis in order to define the molecular basis of B-cell epitope expression.

The high degree of sequence homology of the 3' terminal portion of these cDNA clones with maize mitochondrial 26S rRNA is surprising, but it is confirmed by the hybridization of radiolabelled subfragments of these inserts in Northern blot analysis (Fig. 3). It is possible that these inserts represent highly unusual recombination events occurring during initial cDNA preparation with mRNA containing substantial amounts of copurified rRNA, and do not therefore represent true *Dac g II* transcripts. The loss of the 5' *EcoRI* site used for initial insertion of the DNA and the lack of a polyadenylation signal in the 3' half of the transcript provide circumstantial evidence that these recombination events occurred during cloning into λ gt11, since there is no evidence of an internal *EcoRI* site at or near the junction between the *Dac g II* and the rRNA sequence (Fig. 2a). It is also noteworthy that these cDNA clones were isolated from a library generated from anther, rather than pollen, mRNA (33), and there are differences in expression of ribosomal RNA transcripts between pollen and anther.

Northern blot probing with intact inserts and subfragments suggests that the 5' end of these *Dac g II* transcripts may be incomplete, since an approximately 800-bp transcript is detected in *Dac g* pollen RNA (Fig. 3c). That this band (or comigrating transcripts) is also detected by the full-length and 3' region of the allergen clones used in this study may

be interpreted as supporting both views of the origin of these inserts; that is, that they arose by recombination events involving copurified ribosomal RNA transcripts, or that they represent original transcripts which have portions of the noncoding sequence bearing homology to ribosomal RNA sequences.

Combined with the sequence data, the Northern blot studies do, however, suggest that group II, and possibly group III, grass pollen allergens are encoded by a gene or gene family distinct from that encoding major (groups I and V/IX) allergens, and that these allergens are differentially expressed in anther and pollen with differing *in vivo* functions. Similar differences in cellular function and tissue expression have been previously demonstrated for *Lol p I* and *Lol p V/IX* (31).

Acknowledgments

The authors would like to thank Drs David Walsh and Irwin Griffith for useful comments and advice during these studies. This work was supported by grants from the Medical Research Council, UK, the National Asthma Campaign, UK, and the Wellcome Trust, UK.

References

1. ANSARI AA, KIHARA TK, MARSH DG. Immunochemical studies of *Lolium perenne* (rye grass) pollen allergens, *Lol p I*, II, and III. *J Immunol* 1987; 139: 4034–41.
2. ANSARI AA, SHENBAGAMURTHI P, MARSH DG. Complete primary structure of a *Lolium perenne* (perennial rye grass) pollen allergen, *Lol p II*. *J Biol Chem* 1989; 264: 11181–5.
3. ANSARI AA, SHENBAGAMURTHI P, MARSH DG. Complete primary structure of a *Lolium perenne* (perennial rye grass) pollen allergen, *Lol p III*: comparison with known *Lol p I* and II sequences. *Biochemistry* 1989; 28: 8665–70.
4. DEVERAUX J, HAEGERLI P, SMITHIES O. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 1984; 12: 387–95.
5. ESCH RE, KLAPPER DG. Cross-reactive and unique grass group I antigenic determinants defined by monoclonal antibodies. *J Allergy Clin Immunol* 1987; 79: 489–95.
6. ESCH RE, KLAPPER DG. Isolation and characterisation of a major cross-reactive grass group I allergenic determinant. *Mol Immunol* 1989; 26: 557–61.
7. FEINBERG AP, VOGELSTEIN B. A technique for radiolabeling DNA restriction fragments to a high specific activity. *Anal Biochem* 1984; 137: 266–7.
8. FREIDHOFF LR, KAUTZKY EE, GRANT JH, MEYERS D, MARSH DG. A study of the human immune response to *Lolium perenne* (rye) pollen and its components, *Lol p I* and *Lol p II* (rye I and rye II). I. Prevalence of reactivity to the allergens and correlations among skin test, IgE antibody and IgG antibody data. *J Allergy Clin Immunol* 1986; 78: 1190–1201.
9. FREIDHOFF LR, KAUTZKY EE, MEYERS DA, ANSARI AA, BIAS WB, MARSH DG. Association of HLA-DR3 and total serum immunoglobulin E level with human immune response to *Lol p I* and *Lol p II* allergens in allergic subjects. *Tissue Antigens* 1988; 31: 211–19.
10. GERLACH WL, BEDBROOK JR. Cloning and characterisation of ribosomal RNA genes from wheat and barley. *Nucleic Acids Res* 1979; 7: 1869–85.
11. GRIFFITH TJ, SMITH PM, POLLOCK J, et al. Cloning and sequencing of *Lol p Ia*, the major allergenic fraction of rye grass pollen. *FEBS Lett* 1991; 279: 210–15.
12. HANSON DD, HAMILTON DA, TRAVIS JL, BASHE DM, MASCARENHAS JP. Characterization of a pollen specific cDNA clone from *Zea mays* and its expression. *Plant Cell* 1989; 1: 173–9.
13. JEPSON I, BRAY J, JENKINS G, SCHUCH W, EDWARDS K. A rapid procedure for the construction of PCR cDNA libraries from small amounts of plant tissue. *Plant Mol Biol Rep* 1991; 9: 131–8.
14. LAEMMLI UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970; 227: 680–2.
15. LIN Z, EKRAMODDOULLAH AKM, KISIL FT, HERBERT J, MOURAD W. Isolation and characterization of *Poa p I* allergens of Kentucky bluegrass pollen with a murine monoclonal *Lol p I* antibody. *Int Arch Allergy Appl Immunol* 1988; 87: 294–300.
16. MARSH DG, GOODFRIEND L, KING TP, LÖWENSTEIN H, PLATTS-MILLS TAE. Allergen nomenclature. *J Allergy Clin Immunol* 1987; 80: 639–45.
17. MATTHIESEN F, LÖWENSTEIN H. Group V allergens in grass pollen. I. Purification and characterisation of the group V allergen from *Phleum pratense* pollen, *Phl p V*. *Clin Exp Allergy* 1991; 21: 297–307.
18. MATTHIESEN F, LÖWENSTEIN H. Group V allergens in grass pollen. II. Identification of group V allergens in pollens from eight temperate grasses. *Clin Exp Allergy* 1991; 21: 309–20.
19. MERCHERI S, PELTRE G, DAVID B. Purification and characterisation of a major allergen from *Dactylis glomerata* pollen: the AgDgl. *Int Arch Allergy Appl Immunol* 1985; 78: 283–9.
20. MOHAPATRA SS, HILL R, ASTWOOD J, et al. Isolation and characterisation of a cDNA clones encoding and IgE-binding protein from Kentucky bluegrass (*Poa pratense*) pollen. *Int Arch Allergy Appl Immunol* 1990; 91: 362–8.
21. MOURAD W, MECHERI S, PELTRE G, DAVID B, HERBERT J. Study of the epitope structure of purified *Dac g I* and *Lol p I* the major allergen of *Dactylis glomerata* and *Lolium perenne* pollens using monoclonal antibodies. *J Immunol* 1988; 141: 3486–91.
22. PEREZ M, ISHIOKA GY, WALKER LE, CHESNUT RW. cDNA cloning and immunological characterisation of the rye grass allergen I. *J Biol Chem* 1990; 265: 16210–15.
23. RAPLEY R, FLORA PF, WALKER MR. Direct PCR sequencing of murine immunoglobulin genes using *E. coli* single stranded binding protein. *PCR Methods Appl* 1992; 2: 99–101.
24. VAN REE R, CLEMENS JGJ, AALBERS M, STAPEL SO, AALBERSE RC. Characterisation with monoclonal and polyclonal antibodies of a new major allergen from grass pollen in the group I molecular weight range. *J Allergy Clin Immunol* 1989; 83: 144–51.
25. ROBERTS AM, VAN REE R, CARDY SM, BEVAN LJ, WALKER MR. Recombinant pollen allergens from *Dactylis glomerata*: preliminary evidence that human IgE cross-reacting with *Dac g II* and *Lol p I/II* is increased following grass pollen immunotherapy. *Immunology* 1992; 76: 389–96.
26. ROCKLIN RE. Clinical and immunologic aspects of allergen-specific immunotherapy in patients with seasonal allergic rhinitis and/or allergic asthma. *J Allergy Clin Immunol* 1988; 72: 323–9.
27. SAMBROOK J, FRISCHT EF, MANIATIS T. Molecular cloning: a laboratory manual. 2nd ed. New York: Cold Spring Harbor, 1990.
28. SARKAR G, SOMMER SS. Access to a messenger RNA sequence or its protein product is not limited by tissue or species specificity. *Science* 1989; 244: 331–4.

29. SOMMER SS, SARKAR G, KOEBERL DD, et al. Direct sequencing with the aid of phage promoters. In: INNIS MA, GELFAND DH, SNINSKY JJ, WHITE TJ, eds. PCR protocols: a guide to methods and applications. San Diego, CA: Academic Press, 1990: 197–205.
30. SILVANOVIČ 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.
31. SINGH MB, HOUGH T, THEERAKULPISUT P, et al. Isolation of complementary DNA encoding a newly identified major allergenic protein of rye-grass pollen: intracellular targeting to the amyloplast. Proc Natl Acad Sci U S A 1991; 88: 1384–8.
32. THOMAS WR. Grass under scrutiny. Clin Exp Allergy 1991; 21: 255–7.
33. WALSH DJ, MATTHEWS JA, DENMEADE R, WALKER MR. Cloning of cDNA coding for an allergen of cocksfoot grass (*Dactylis glomerata*) pollen. Int Arch Allergy Appl Immunol 1989; 90: 78–83.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.