

Pollen allergens are restricted to few protein families and show distinct patterns of species distribution

Christian Radauer, PhD, and Heimo Breiteneder, PhD *Vienna, Austria*

Background: Inhalative allergies are elicited predominantly by pollen of various plant species. However, a classification of the large number of identified pollen allergens is still missing.

Objective: To analyze pollen allergen sequences with respect to protein family membership, taxonomic distribution of protein families, and interspecies variability.

Methods: Protein family memberships of all plant allergen sequences from the Allergome database were determined by using the Protein Families Database of Alignments and Hidden Markov Models. The taxonomic distribution of pollen allergens was established from the Integrated Taxonomic Information System. Members of abundant pollen allergen families were compared with allergenic and nonallergenic homologues by database similarity searches and multiple sequence alignments. **Results:** Pollen allergens were classified into 29 of 7868 protein families. Expansins, profilins, and calcium-binding proteins constitute the major pollen allergen families, whereas most plant food allergens belong to the prolamin, cupin, or profilin families. Pollen allergens were revealed to be ubiquitous (eg, profilins), present in certain plant families (eg, pectate lyases), or limited to a single taxon (eg, thaumatin-like proteins). Allergenic plant profilins constitute a highly conserved family with sequence identities of 70% to 85% among each other but low identities of 30% to 40% with nonallergenic profilins from other eukaryotes, including human beings. Similarly, allergenic polcalcins possess sequence identities of 64% to 92% but show low identities of 39% to 42% to related nonallergenic calmodulins and calmodulin-like proteins from vegetative plant tissues and man.

Conclusion: This classification of pollen allergens into protein families will aid in predicting cross-reactivity, designing comprehensive diagnostic devices, and assessing the allergenic potential of novel proteins. (*J Allergy Clin Immunol* 2006;117:141-7.)

Key words: Pollen allergens, protein families, taxonomy, cross-reactivity

Abbreviations used

BLAST: Basic Local Alignment Search Tool
Pfam: Protein Families Database of Alignments and Hidden Markov Models
Uniprot: Universal Protein Resource

Pollen from wind-pollinated seed plants constitutes one of the most important elicitors of type 1 allergy worldwide.¹⁻⁶ A large number of grass, tree, and weed species sheds their pollen in high concentrations during the pollen season, leading to allergic symptoms from seasonal rhinoconjunctivitis to bronchial asthma in susceptible individuals.

The last 2 decades were marked by large advances in the molecular characterization of pollen-derived allergenic proteins. Pollen allergens belonging to various protein families have been identified in diverse plant species.⁷⁻¹⁰ A large amount of information on pollen allergens can be found in allergen databases such as the Allergome (www.allergome.org) or the Official List of Allergens from the International Union of Immunological Societies Allergen Nomenclature Sub-Committee (www.allergen.org). However, a systematic—taxonomic and molecular—classification of pollen allergens has not been made available so far. Such a compilation of taxonomic and molecular relationships among pollen allergens will be a valuable tool for several areas of clinical practice, basic research, and regulatory authorities.

Pure natural or recombinant allergens are currently introduced into diagnostic tests¹¹ and will probably be also used for allergen-specific immunotherapy.¹² However, accuracy of diagnostic tests and efficacy of therapeutic approaches are often compromised by allergenic cross-reactivity both among different pollen species¹³ as well as between pollen and plant-derived food.¹⁴ A systematic molecular classification of pollen allergens will facilitate theoretical understanding and enable prediction of cross-reactivity.

Environmental changes brought about by human civilization may also have considerable effects on pollen allergy. Global climate change affects start, duration, and intensity of pollen seasons, distribution of plant species, and allergen content of pollen grains.¹⁵ Likewise, air pollution may also have effects on the interaction between

From the Department of Pathophysiology, Center of Physiology and Pathophysiology, Medical University of Vienna.

Supported by grant SFB-F01802 from the Austrian Science Fund.

Received for publication May 24, 2005; revised September 13, 2005; accepted for publication September 14, 2005.

Available online November 29, 2005.

Reprint requests: Heimo Breiteneder, PhD, Department of Pathophysiology, Center of Physiology and Pathophysiology, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria. E-mail: Heimo.Breiteneder@meduniwien.ac.at.

0091-6749/\$32.00

© 2005 American Academy of Allergy, Asthma and Immunology

doi:10.1016/j.jaci.2005.09.010

allergens and the human immune system.¹⁶ Finally, human activities have led to the introduction of allergenic non-endemic plants, either accidentally, such as the spread of ragweed in Europe,¹⁷ or by planting foreign plants in urban areas.¹⁸

Thus, in a changing environment a deeper understanding of factors that govern allergenicity of pollen and make specific pollen proteins allergenic will be required for the development of rational strategies for prevention and management of pollen allergy. The aim of this study was, therefore, an analysis of the Allergome database, currently the most comprehensive collection of allergen data, with respect to taxonomic distribution of pollen allergen sources and identified pollen allergens as well as a molecular classification of pollen allergens by alignment of their sequences with the Protein Families Database of Alignments and Hidden Markov Models (Pfam).

METHODS

Allergen sequences and protein family assignment

A list of all pollen and plant food allergens (single sequence per allergen) was extracted from the Allergome database (<http://www.allergome.org>, release of March 1, 2005).¹⁹ Protein sequences of these allergens were obtained from the Universal Protein Resource (Uniprot) database release 4.2 (March 1, 2005) located at the European Bioinformatics Institute (<http://www.ebi.ac.uk>).²⁰ Protein family memberships of each allergen were assigned by comparing the sequences with the Pfam database version 17.0 (<http://www.sanger.ac.uk/Software/Pfam/>, March 2005).²¹ Multidomain allergens were included once for each unique domain if different protein architectures that contained one of these domains were found.

Fragment sequences, such as those obtained by translation of expressed sequence tags or N-terminal protein sequencing, that produced no significant alignment with the Pfam database were assigned to a protein family by comparing them with Viridiplantae sequences from the Uniprot database using the Basic Local Alignment Search Tool (BLAST)²² located at the Swiss Institute of Bioinformatics (<http://www.expasy.org/tools/blast/>) with parameters optimized for short query sequences (comparison matrix, PAM30; E-value threshold, 100, low-complexity filter off, ungapped alignment for fragments shorter than 30 residues). If all of the most significant hits belonged to the same protein family, the allergen in question was assigned to that family.

Protein family distribution of all seed plant protein sequences

To assess whether the distributions of protein families among pollen and plant food allergens reflect the frequencies of these families among all seed plant proteins, all Spermatophyta protein sequence entries containing a Pfam annotation were downloaded from the Uniprot database. The number of occurrence of Pfam domains was counted, with proteins containing several different domains being counted once for each unique domain.

Taxonomic distribution of pollen allergen sources

A list of plant species whose pollen was described to elicit allergic reactions or bind human IgE was obtained from the Allergome database. Botanical taxonomic data were retrieved (March 15, 2005) from the Integrated Taxonomic Information System online database (<http://www.itis.usda.gov>).

Sequence alignments of allergen sequences with their homologues

Representative members of the most abundant pollen allergen families were selected as reference sequences and compared with the Uniprot database using the BLAST server located at the Swiss Institute of Bioinformatics. Sequences of allergenic as well as non-allergenic proteins with significant similarities to the reference sequences (E-value below 10) were classified by protein family and biological source. Ranges of sequence identities to the reference allergen sequence within each group were determined by using the alignments generated by the BLAST program.

Clustering of pollen allergen sequences with allergenic and nonallergenic homologues

To visualize relationships within subfamilies of allergenic and non-allergenic members of the most abundant pollen allergen families, homologues of profilins, expansins, and calcium-binding proteins were subjected to a cluster analysis. A representative sample of sequences from each group defined in the BLAST search described was downloaded and aligned by using the ClustalX 1.83 program.²³ Unrooted neighbor-joining trees generated from these alignments were drawn by using TreeView 1.6.6.²⁴

RESULTS

Protein families of pollen allergens

At the date of data retrieval for this study (March 1, 2005), the Allergome database listed 157 completely or partially sequenced pollen allergens. Comparison of their sequences with the Pfam protein family database revealed a limited representation of protein families among pollen allergens (Fig 1, A). Although members of 2615 protein families (of a total of 7868 protein families in the Pfam database) were found among all seed plant proteins, only 29 protein families contained pollen allergens. The distribution of protein families among pollen allergens differed from the distribution of food allergen families (Fig 1, B; Table I). Sequences of 137 plant food allergens were classified into 27 protein families. Ten of 29 pollen allergen families were also represented among food allergens. Pollen allergen families with at least 3 members are listed in Table I.

The distributions of pollen and plant food allergen families did not correspond to the frequencies of protein families among all seed plant proteins (Fig 1, C; Table I). Among the 25 most abundant families of seed plant proteins, only a single family contained allergens: the prolamin family, which was the most abundant family of plant food allergens but ranked 12th (of 29 protein families) among pollen allergens and 23rd (of 2615 protein families) among all seed plant proteins.

Taxonomic distribution of pollen allergens

The Allergome database listed 178 plant species that contained pollen to which allergic reactions have been described in the literature (see this article's Table E1 in the Online Repository at www.jacionline.org). These plants belonged to 44 taxonomic families and 29 orders. Only 6 of these families contained more than half of the species. About ¼ of allergenic pollen species were grass pollen (Poaceae, 43 species); other abundant families were

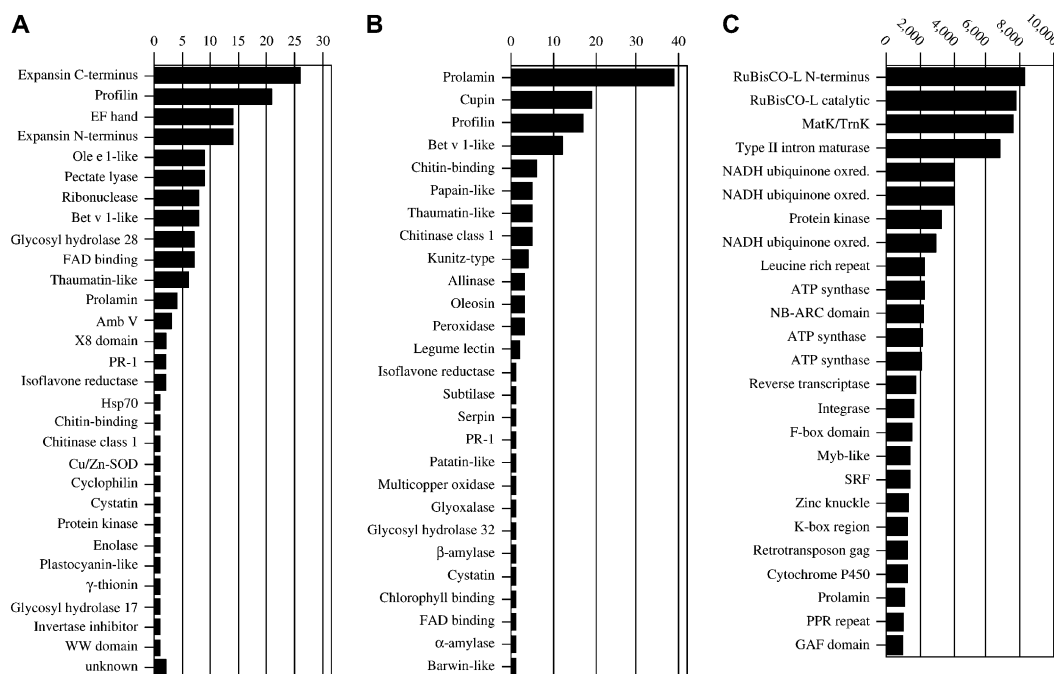


FIG 1. Protein family distributions of pollen and plant food allergens. Allergens from the Allergome database were compared with the Pfam protein family database. Numbers of sequences in all families of pollen (**A**, 29 families) and plant food (**B**, 27 families) allergens and in the 25 most abundant families of all seed plant proteins (**C**, 2615 families total) are shown. *ATP synthase*, adenosine triphosphate synthase; *Cu/Zn-SOD*, copper-zinc superoxide dismutase; *FAD binding*, flavin adenine dinucleotide binding protein; *Hsp70*, heat shock protein 70; *NADH ubiquinone oxred.*, nicotinamide adenine dinucleotide ubiquinone oxidoreductase; *PPR repeat*, pentatricopeptide repeat; *PR-1*, pathogenesis related protein 1; *RuBisCo-L*, ribulose biphosphate carboxylase-oxygenase large subunit; *SRF*, serum response factor-type transcription factor.

TABLE I. Distribution of pollen allergens into protein families and comparison with their occurrence within plant food allergens and all seed plant proteins*

Family	Pfam accession	Pollen allergens		Plant food allergens		All seed plant proteins	
		Rank (of 29)	Number† (of 157)	Rank (of 27)	Number† (of 137)	Rank (of 2615)	Number (of 170,661)
Expansin C-terminus	PF01357	1	26	—	—	53	402
Profilin	PF00235	2	21	3	17	279	94
EF-hand	PF00036	3	14	—	—	34	606
Expansin N-terminus	PF03330	3	14	—	—	68	347
Ole e 1-like	PF01190	5	9	—	—	347	73
Pectate lyase	PF00544	5	9	—	—	317	82
Ribonuclease	PF01620	7	8	—	—	507	45
Bet v 1-like	PF00407	7	8	4	12	59	382
Glycosyl hydrolase 28	PF00295	9	7	—	—	116	226
FAD-binding domain	PF01565	9	7	14	1	208	140
Thaumatococcus-like	PF00314	11	6	6	5	105	258
Prolamin	PF00234	12	4	1	39	23	1075
Amb V allergens	PF03913	13	3	—	—	1921	4

*Only pollen allergen families with at least 3 members are listed.

†Single sequence per allergen.

Cupressaceae (cypress family, 16 species), Betulaceae (birch family, 12 species), Arecaceae (palms, 8 species), Asteraceae (ragweed family, 8 species), and Fabaceae (legumes, 7 species).

In contrast with the large number of allergenic pollen species, the taxonomic repertoire of sequenced pollen

allergens was limited to 52 species from 14 taxonomic families (see this article's [Table E1](#) in the Online Repository at www.jacionline.org). Plant families with the largest number of sequenced allergens were Poaceae (71 allergens from 15 species), Cupressaceae (20 allergens from 8 species), Asteraceae (13 allergens from 5 species),

TABLE II. Taxonomic distribution of pollen allergen families*

Family	Number of allergens												
	Total	Expansin C-term	Profilin	EF hand	Expansin N-term	Ole e 1-like	Pectate lyase	Ribonu-lease	Bet v 1-like	Glyco hydr 28	FAD binding	Thaumatococcal	Amb V
Cupressaceae	20			1			7			3		6	
Arecaceae	1		1										
Poaceae	71	26	6	2	14	3		8		2	7		
Asteraceae	13		3				2						3
Plantaginaceae	1					1							
Oleaceae	12		1	3		4							
Chenopodiaceae	4		1	1		1							
Brassicaceae	8		2	4						1		1	
Betulaceae	13		2	3					5				
Fagaceae	3								3				
Platanaceae	2									1			
Cannabaceae	2		1										
Urticaceae	3		1										2
Euphorbiaceae	2		2										

C-term, C-terminal domain; N-term, N-terminal domain; Glyco hydr 28, glycosyl hydrolase family 28; FAD binding, flavin adenine dinucleotide binding protein.

*Only pollen allergen families with at least 3 members are listed.

Betulaceae (13 allergens from 5 species), and Oleaceae (12 allergens from 4 species).

An arrangement of pollen allergens by protein family and taxonomic family (Table II) reveals large differences in the taxonomic distribution of pollen allergen families. Ubiquitous pollen allergen families occurring in almost all plant families examined include the profilins and the EF hand containing calcium-binding proteins. Several allergen families occurred in a small number of taxonomically diverse plant families, such as the Ole e 1 homologues, pectate lyases, polygalacturonases (members of the glycosyl hydrolase family 28), and nonspecific lipid transfer proteins (members of the prolamin superfamily). The third group consisted of protein families confined to single taxa. Allergenic expansins, ribonucleases, and flavin adenine dinucleotide-binding proteins were confined to grass (Poaceae) pollen, Bet v 1 homologous allergens were specific for pollen from the order Fagales (Betulaceae and Fagaceae), and the thaumatin-like and the Amb V allergens were found exclusively in Cupressaceae and Asteraceae pollen, respectively.

Interspecies variability of pollen allergens and their homologues

Sequence similarities among members of prevailing pollen allergen families and their allergenic and nonallergenic homologues are depicted in Fig 2. In addition, multiple sequence alignments of members of the 3 most abundant pollen allergen families—profilins, expansins, and calcium-binding proteins—and their homologues were used to generate trees representing divergence between any pair of sequences (Fig 3).

The comparison of plant profilins with homologues from fungi and vertebrates (Fig 3, A) revealed that they form a highly conserved family with sequence identities between 70% and 85% among each other, but low identities (30% to 40%) compared with nonallergenic profilins

from other eukaryotes, including human beings (Figs 2, 3, A).

The C-terminal expansin domain was found within 2 groups of grass pollen allergens, the group 1 allergens that also contained the N-terminal expansin-domain and the group 2/3 allergens lacking this part of the sequence. Group 1 grass pollen allergens shared 41% to 90% of their sequences, whereas group 2/3 allergens showed 45% to 46% sequence identity to the reference sequence of Phl p 1 (Fig 2). Both allergen families were related to β -expansins from dicotyledonous plants (Fig 3, B). However, sequence similarities to nonallergenic dicot β -expansins as well as to α -expansins from monocots and dicots were equally low (27% to 47%).

The largest group of pollen allergens containing EF hand domains is the polcalcin family, which is characterized by the presence of 2 EF hand domains and sequence identities of 64% to 92% (Figs 2, 3, C). A small number of pollen allergens with 3 (Bet v 3 from birch) or 4 (Ole e 8 from olive and Jun o 2 from juniper) of these structural elements has also been characterized and displayed 51% to 59% sequence identities to polcalcins (Figs 2, 3, C). All of these allergens showed low sequence similarities to calmodulins and calmodulin-like proteins from vegetative plant tissue as well as from animals, including human beings (39% to 42%).

Members of the Ole e 1 family were conserved among Oleaceae species (86% to 89% identity), but showed low similarities of 32% to 39% to homologues from other families (Fig 2). Similarly, among members of the pectate lyase family, sequences of ragweed group 1/2 allergens, Cupressaceae group 1 allergens, and nonallergenic pectate lyases from fruits diverged considerably (45% to 49% identity; Fig 2). Sequences of group 5/6 grass pollen allergens were quite variable (48% to 88% identity to Phl p 5). Bet v 1 homologues from Fagales tree pollen showed sequence identities above 70% among each other, but

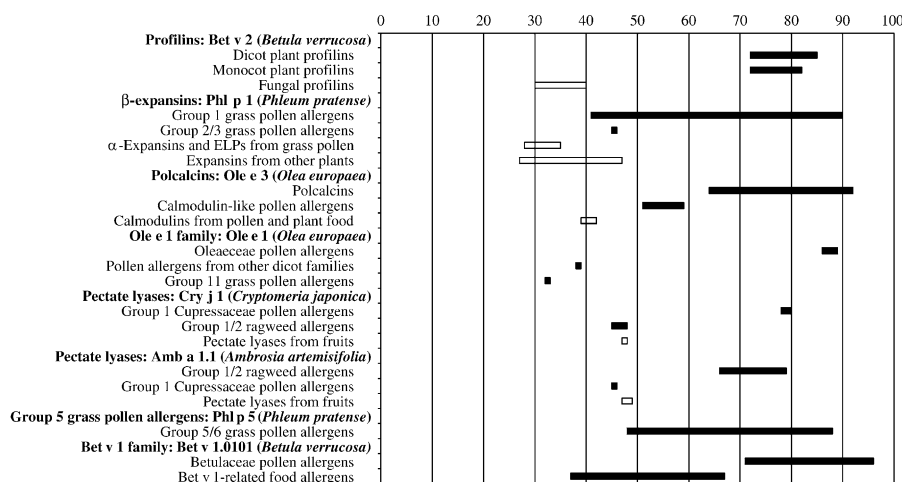


FIG 2. Sequence variability among important pollen allergen families and their allergenic and nonallergenic homologues. Representative allergens from each family (**bold**) were compared with the Uniprot database by using the BLAST program. The diagram shows ranges of percent sequence identity of allergenic (**black bars**) and nonallergenic (**white bars**) homologues. ELPs, expansin-like proteins.

much lower similarities to homologues from typical birch pollinosis-related fruits and vegetables (37% to 67%; Fig 2).

DISCUSSION

Many pollen allergens are among the best characterized allergenic proteins studied so far. The cloning of the first pollen allergen, the major birch pollen allergen, Bet v 1, in 1989²⁵ marked the advent of a decade characterized by large progresses in the molecular characterization of allergens. Despite the wealth of data compiled until now, a systematic classification of pollen allergens by protein family and taxonomic origin has been missing.

Comparison of 157 at least partially sequenced pollen allergens with the Pfam protein family database revealed that pollen allergens were distributed among only 29 protein families (Fig 1, A), a small number compared with a total of 2615 protein families (of 7868 families in the Pfam database) found in seed plants. An equally limited protein family distribution was found among plant food allergens (Fig 1, B). Both distributions corroborate data from a recent publication based on different allergen databases and an earlier Pfam version.²⁶ In a study aimed at implementing an allergy prediction system based on sequence motifs, Stadler and Stadler²⁷ found only 52 motifs that matched 644 of 779 allergen sequences from all types of sources. Thus, limited occurrence of allergens within protein families is not confined to allergens of plant origin.

The patterns of protein families containing pollen allergens showed marked differences from the distribution of plant food allergens (Table I). Whereas expansins, profilins, and calcium-binding proteins constitute the major pollen allergen families, most plant food allergens belong to the prolamin, cupin, and profilin families. Families important in both types of allergen sources are the profilins and the Bet v 1 homologues, with both families

responsible for the well known pollen-associated food allergies.¹⁴ Several allergen families have been described both in pollen and plant food: the flavin adenine dinucleotide-binding proteins,²⁸ thaumatin-like proteins,²⁹ lipid-transfer proteins (members of the prolamin superfamily),³⁰ and the isoflavone reductases.³¹ However, data on their cross-reactivity between pollen and food are scarce.

The spectrum of plant species containing allergenic pollen is much larger than the number of species with allergens studied on the molecular level (see this article's Table E1 in the Online Repository at www.jacionline.org). A comparison of the taxonomic distribution of sequenced allergens and all potentially allergenic pollen species revealed 2 types of discrepancies. The first bias is a taxonomic one. The group of plants studied in most detail are grasses, which are confined to a single taxonomic family, Poaceae. Seventy-one allergens from 15 species, distributed among 10 protein families, have already been isolated. Considering this together with the high cross-reactivity among grass pollen species,⁷ it can be assumed that most relevant grass pollen allergens have already been identified. Trees are a taxonomically more diverse group. Members of several families such as Cupressaceae,⁸ Betulaceae,¹⁰ and Oleaceae³² have been extensively studied, whereas characterization of allergens from other common trees such as palms (Arecaceae), maples (Aceraceae), or acacias (Fabaceae) is still in its infancy.¹⁰ Identification of weed pollen allergens, although taxonomically most diverse, has concentrated so far on mugwort and several species of ragweed and pellitory.^{9,33} In contrast, members of several plant families containing weed species common in tropical and subtropical climate zones, such as Amaranthaceae and Euphorbiaceae, are poorly characterized.

The second discrepancy between potentially allergenic pollen species and species studied on the molecular level is a geographic one. Whereas wind-pollinated species of Europe, North America, and Japan have been studied in

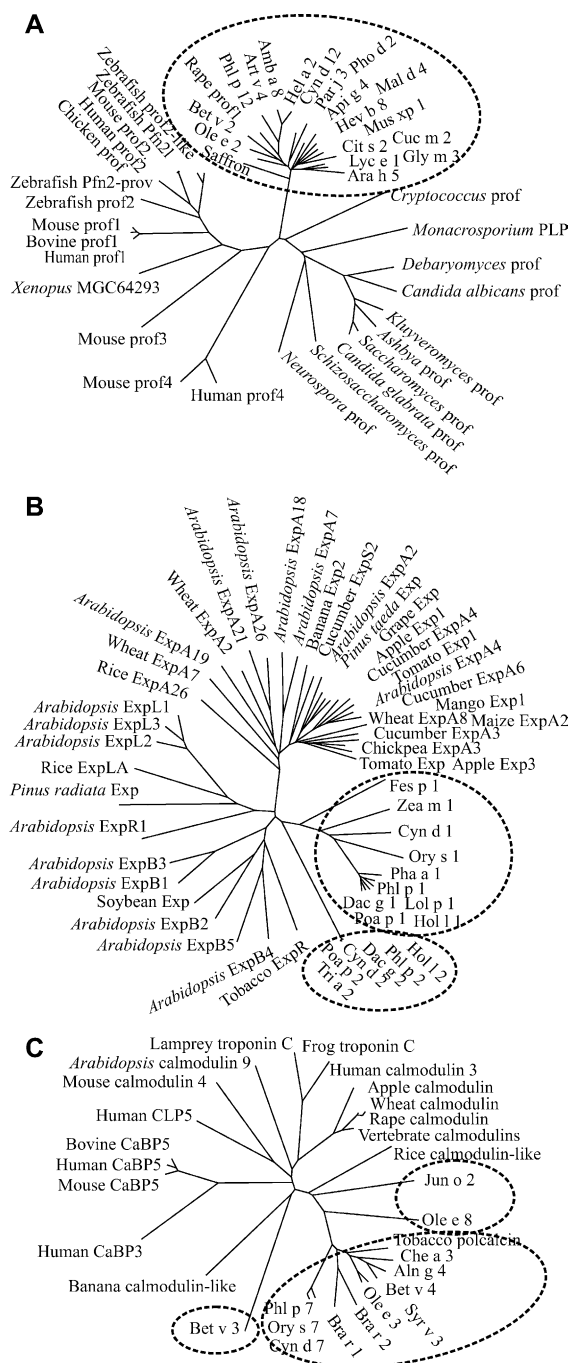


FIG 3. Unrooted neighbor-joining trees generated from sequence alignments of major pollen allergen families and nonallergenic proteins with significant sequence similarities to these allergens. **A**, Profilins; **B**, expansins; **C**, EF-hand proteins. Branch lengths are proportional to sequence divergence. Allergens are marked by dashed ellipses. CaBP, calcium-binding protein; CLP, calmodulin-like protein; Exp, expansin; ExpA, alpha-expansin; ExpB, beta-expansin; ExpL, expansin-like protein; ExpR, expansin-related protein; prof, profilin; PLP, profilin-like protein.

great detail,^{1,5} much less is known about allergens from tropical and subtropical pollen species. However, the prevalence of allergic diseases has been increasing in

developing countries.^{4,34,35} The most striking example of an allergologically important but poorly characterized plant family is the family Arecaceae (palm trees), which contains several species described as major elicitors of pollen allergy in Southern Asia.^{2,3} Opposed to its allergological significance, only a single pollen allergen from the palm family has been cloned so far, Pho d 2, the profilin from date palm (*Phoenix dactylifera*).³⁶

An arrangement of sequenced pollen allergens by protein family and taxonomic family revealed 3 types of allergens (Table II): ubiquitous allergens such as profilins and EF-hand proteins, allergen families restricted to a small number of taxonomically diverse plant families such as Ole e 1 homologues and pectate lyases, and protein families confined to single plant families or orders. Most of these plant family-specific allergen families are found in grass pollen. This may reflect the large number of grass pollen allergens sequenced so far, possibly including allergens present but not detected yet in other pollen species. Another explanation may be the evolutionary distance of grasses from other allergenic plants that may have led to the development of grass-specific pollen proteins.

Interestingly, the prevalence of sensitization to the ubiquitous allergens, profilins and calcium-binding proteins, is usually low. One example is grass pollen. Most patients are sensitized to β -expansins (group 1) and ribonucleases (group 5), with both protein families specific for grass pollen. On the other hand, frequencies of sensitization to profilins (group 12) and calcium-binding proteins (group 7) are below 20%.⁷ Similarly, homologues of the major birch pollen allergen, Bet v 1, recognized by IgE from more than 90% of patients allergic to birch pollen, are found exclusively in trees from the order Fagales. In contrast, profilin (Bet v 2) and calcium-binding proteins (Bet v 3 and Bet v 4) are minor allergens with frequencies of sensitization below 20%.³⁷ One reason for the low allergenic potential of profilins and calcium-binding proteins may be the fact that members of both protein families are found in human beings (Fig 3), leading to suppression of the immune response to these allergens in most individuals.

A comparison of sequence conservation within allergen family members (Fig 2), taxonomic distribution of identified allergens (Table II), and cross-reactivity data from the literature¹³ reveals a high degree of correspondence. Ubiquitous, highly cross-reactive allergens such as profilins and calcium-binding proteins show sequence similarities of more than 60% between members of different plant families. In contrast, β -expansins from grasses share less than 50% of their sequences with homologues from other plant families, which do not cross-react with grass pollen expansins and have never been described as allergens. Similarly, the major olive pollen allergen, Ole e 1, cross-reacts with its homologues from the family Oleaceae (Fra e 1 from ash, Lig v 1 from privet, and Syr v 1 from lilac), which share almost 90% of their sequences.³⁸ In contrast, cross-reactivity to homologues from other plant families (Pla l 1 from plantain and Che a 2 from goosefoot) is weak, which matches their low sequence similarities to Ole e 1.^{39,40} A review of sequence similarities among

allergenic and nonallergenic homologues of pollen allergens (Fig 2) leads to the conclusion that the prerequisite for allergenic cross-reactivity is sequence identity of at least 50%. The only known exception from this rule is the Bet v 1 family. Cross-reactive Bet v 1 homologues that are responsible for birch pollinosis-related food allergy share only 37% to 67% of their sequences with Bet v 1. This may be explained by the fact that surface-exposed residues of Bet v 1 homologues are more conserved than buried ones, resulting in higher levels of similarity between surfaces compared with entire sequences.²⁶

The data obtained in this study give rise to several suggestions for future focuses of pollen allergen research. First, current knowledge of pollen allergens is concentrated on major allergen sources of Europe, North America, and Japan. The increase of allergic diseases, especially in Asia, emphasizes the need for molecular characterization of tropical and subtropical pollen species. Second, even in Western countries, there are several poorly characterized pollen allergen sources—for example, those plants hitherto overlooked because their pollination periods overlap with major elicitors of pollen allergy. Finally, sequence comparisons of allergens with their nonallergenic homologues bring about new stimuli in the elucidation of factors that confer allergenicity on proteins, as exemplified by the apparent lack of allergenicity of α -expansins, fungal profilins, or pectate lyases from fruits.

REFERENCES

- D'Amato G, Spiekma FT, Liccardi G, Jager S, Russo M, Kontou-Fili K, et al. Pollen-related allergy in Europe. *Allergy* 1998;53:567-78.
- Boral D, Chatterjee S, Bhattacharya K. The occurrence and allergising potential of airborne pollen in West Bengal, India. *Ann Agric Environ Med* 2004;11:45-52.
- Chew FT, Lim SH, Shang HS, Dahlia MD, Goh DY, Lee BW, et al. Evaluation of the allergenicity of tropical pollen and airborne spores in Singapore. *Allergy* 2000;55:340-7.
- Sibanda EN. Inhalant allergies in Zimbabwe: a common problem. *Int Arch Allergy Immunol* 2003;130:2-9.
- White JF, Bernstein DI. Key pollen allergens in North America. *Ann Allergy Asthma Immunol* 2003;91:425-35.
- Kaneko Y, Motohashi Y, Nakamura H, Endo T, Eboshida A. Increasing prevalence of Japanese cedar pollinosis: a meta-regression analysis. *Int Arch Allergy Immunol* 2005;136:365-71.
- Andersson K, Lidholm J. Characteristics and immunobiology of grass pollen allergens. *Int Arch Allergy Immunol* 2003;130:87-107.
- Di Felice G, Barletta B, Tinghino R, Pini C. Cupressaceae pollinosis: identification, purification and cloning of relevant allergens. *Int Arch Allergy Immunol* 2001;125:280-9.
- Gadermaier G, Dedic A, Obermeyer G, Frank S, Himly M, Ferreira F. Biology of weed pollen allergens. *Curr Allergy Asthma Rep* 2004;4:391-400.
- Mothes N, Horak F, Valenta R. Transition from a botanical to a molecular classification in tree pollen allergy: implications for diagnosis and therapy. *Int Arch Allergy Immunol* 2004;135:357-73.
- Hamilton RG, Franklin Adkinson N Jr. In vitro assays for the diagnosis of IgE-mediated disorders. *J Allergy Clin Immunol* 2004;114:213-25.
- Wheeler AW, Woroniecki SR. Allergy vaccines: new approaches to an old concept. *Expert Opin Biol Ther* 2004;4:1473-81.
- Weber RW. Patterns of pollen cross-allergenicity. *J Allergy Clin Immunol* 2003;112:229-39.
- Vieths S, Scheurer S, Ballmer-Weber B. Current understanding of cross-reactivity of food allergens and pollen. *Ann N Y Acad Sci* 2002;964:47-68.
- Beggs PJ. Impacts of climate change on aeroallergens: past and future. *Clin Exp Allergy* 2004;34:1507-13.
- Bernstein JA, Alexis N, Barnes C, Bernstein IL, Nel A, Peden D, et al. Health effects of air pollution. *J Allergy Clin Immunol* 2004;114:1116-23.
- Rybníček O, Jäger S. Ambrosia (Ragweed) in Europe. *Allergy Clin Immunol Int* 2001;13:60-6.
- Thompson JL, Thompson JE. The urban jungle and allergy. *Immunol Allergy Clin North Am* 2003;23:371-87.
- Mari A, Riccioli D. The Allergome Web site: a database of allergenic molecules: aim, structure, and data of a Web-based resource. *J Allergy Clin Immunol* 2004;113:S301.
- Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, et al. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res* 2004;32:D115-9.
- Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, et al. The Pfam protein families database. *Nucleic Acids Res* 2004;32:D138-41.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389-402.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997;25:4876-82.
- Page RD. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 1996;12:357-8.
- Breiteneder H, Pettenburger K, Bito A, Valenta R, Kraft D, Rumpold H, et al. The gene coding for the major birch pollen allergen Bet v 1, is highly homologous to a pea disease resistance response gene. *EMBO J* 1989;8:1935-8.
- Jenkins JA, Griffiths-Jones S, Shewry PR, Breiteneder H, Mills EN. Structural relatedness of plant food allergens with specific reference to cross-reactive allergens: an in silico analysis. *J Allergy Clin Immunol* 2005;115:163-70.
- Stadler MB, Stadler BM. Allergenicity prediction by protein sequence. *FASEB J* 2003;17:1141-3.
- Bublin M, Radauer C, Wilson IB, Kraft D, Scheiner O, Breiteneder H, et al. Cross-reactive N-glycans of Api g 5, a high molecular weight glycoprotein allergen from celery, are required for immunoglobulin E binding and activation of effector cells from allergic patients. *FASEB J* 2003;17:1697-9.
- Breiteneder H. Thaumatin-like proteins: a new family of pollen and fruit allergens. *Allergy* 2004;59:479-81.
- Salcedo G, Sanchez-Monge R, Diaz-Perales A, Garcia-Casado G, Barber D. Plant non-specific lipid transfer proteins as food and pollen allergens. *Clin Exp Allergy* 2004;34:1336-41.
- Karamloo F, Wangorsch A, Kasahara H, Davin LB, Hausteiner D, Lewis NG, et al. Phenylcoumaran benzylic ether and isoflavonoid reductases are a new class of cross-reactive allergens in birch pollen, fruits and vegetables. *Eur J Biochem* 2001;268:5310-20.
- Rodriguez R, Villalba M, Monsalve RI, Batanero E. The spectrum of olive pollen allergens. *Int Arch Allergy Immunol* 2001;125:185-95.
- Colombo P, Bonura A, Costa M, Izzo V, Passantino R, Locorotondo G, et al. The allergens of Parietaria. *Int Arch Allergy Immunol* 2003;130:173-9.
- Hong SJ, Lee MS, Sohn MH, Shim JY, Han YS, Park KS, et al. Self-reported prevalence and risk factors of asthma among Korean adolescents: 5-year follow-up study, 1995-2000. *Clin Exp Allergy* 2004;34:1556-62.
- Masoli M, Fabian D, Holt S, Beasley R. The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy* 2004;59:469-78.
- Asturias JA, Ibarrola I, Fernandez J, Arilla MC, Gonzalez-Rioja R, Martinez A. Pho d 2, a major allergen from date palm pollen, is a profilin: cloning, sequencing, and immunoglobulin E cross-reactivity with other profilins. *Clin Exp Allergy* 2005;35:374-81.
- Moverare R, Westritschnig K, Svensson M, Hayek B, Bende M, Pauli G, et al. Different IgE reactivity profiles in birch pollen-sensitive patients from six European populations revealed by recombinant allergens: an imprint of local sensitization. *Int Arch Allergy Immunol* 2002;128:325-35.
- Lombardero M, Obispo T, Calabozo B, Lezaun A, Polo F, Barber D. Cross-reactivity between olive and other species. Role of Ole e 1-related proteins. *Allergy* 2002;57(suppl 71):29-34.
- Calabozo B, Diaz-Perales A, Salcedo G, Barber D, Polo F. Cloning and expression of biologically active *Plantago lanceolata* pollen allergen Pla l 1 in the yeast *Pichia pastoris*. *Biochem J* 2003;372:889-96.
- Barderas R, Villalba M, Lombardero M, Rodriguez R. Identification and characterization of Che a 1 allergen from *Chenopodium album* pollen. *Int Arch Allergy Immunol* 2002;127:47-54.

TABLE E1. Taxonomic distribution of pollen allergen sources

Division	Class	Subclass	Order	Family (typical members)	Species*
Coniferophyta	Pinopsida		Pinales	Cupressaceae (cypress, juniper)	16 (8)
				Pinaceae (pine)	6 (0)
				Podocarpaceae (fern pine)	2 (0)
			Taxales	Taxaceae (yew)	1 (0)
Ginkgophyta	Ginkgoopsida		Ginkgoales	Ginkgoaceae (ginkgo)	1 (0)
Magnoliophyta	Liliopsida	Arecidae	Arecales	Arecaceae (palms)	8 (1)
		Commelinidae	Cyperales	Cyperaceae (sedges)	2 (0)
				Poaceae (grasses)	43 (15)
		Zingiberidae	Zingiberales	Cannaceae	1 (0)
	Magnoliopsida	Asteridae	Asterales	Asteraceae (ragweed, mugwort)	8 (5)
			Dipsacales	Adoxaceae (elder)	1 (0)
			Lamiales	Boraginaceae	1 (0)
			Plantaginales	Plantaginaceae (plantain)	1 (1)
			Scrophulariales	Oleaceae (olive, ash)	5 (4)
			Solanales	Convolvulaceae (morning glory)	1 (0)
		Caryophyllidae	Caryophyllales	Amaranthaceae (amaranth)	6 (0)
				Chenopodiaceae (lamb's quarters)	3 (2)
		Dilleniidae	Capparales	Brassicaceae (mustard)	4 (3)
			Ebenales	Sapotaceae (moa tree)	1 (0)
			Malvales	Bombacaceae (cotton tree)	1 (0)
				Tiliaceae (basswood)	1 (0)
			Salicacales	Salicaceae (poplar)	1 (0)
			Violales	Tamaricaceae (tamarisk)	1 (0)
		Hamamelidae	Casuarinales	Casuarinaceae (horsetail tree)	1 (0)
			Fagales	Betulaceae (birch, alder, hazel)	12 (5)
				Fagaceae (beech, oak, chestnut)	4 (3)
			Hamamelidales	Platanaceae (plane tree)	1 (1)
			Juglandales	Juglandaceae (walnut)	1 (0)
			Myricales	Myricaceae (wax myrtle)	1 (0)
			Urticales	Cannabaceae (hemp, hop)	2 (1)
				Moraceae (mulberry)	1 (0)
				Ulmaceae (elm)	3 (0)
				Urticaceae (pellitory)	4 (1)
		Rosidae	Euphorbiales	Euphorbiaceae (annual mercury)	5 (2)
			Fabales	Fabaceae (acacia, cassia)	7 (0)
			Rhamnales	Elaeagnaceae (Russian olive)	1 (0)
			Myrtales	Melastomataceae (glory tree)	1 (0)
				Myrtaceae (eucalyptus)	4 (0)
			Sapindales	Aceraceae (maple)	5 (0)
				Anacardiaceae (pistachio)	6 (0)
				Meliaceae (neem tree)	1 (0)
				Sapindaceae (horse chestnut)	1 (0)
				Simaroubaceae (tree of heaven)	1 (0)
				Zygophyllaceae (beancaper)	1 (0)

*Number of species with allergenic pollen and with sequenced allergens, respectively.