

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

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

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

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

Download English Version:

<https://daneshyari.com/en/article/3203848>

Download Persian Version:

<https://daneshyari.com/article/3203848>

[Daneshyari.com](https://daneshyari.com)