

Evolutionary biology of plant food allergens

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

The majority of plant food allergens can be grouped into just 4 protein families. This review summarizes the evolutionary relationships of allergenic and nonallergenic members of these families. Proteins from the prolamin superfamily have been described in vascular plants. This superfamily contains several allergenic (2S albumins, nonspecific lipid transfer proteins, and cereal amylase and protease inhibitors) and nonallergenic (hybrid proline-rich proteins, cereal indolines, and α -globulins) member families. The cupin superfamily comprises numerous functionally highly diverse protein families from all groups of organisms. However, allergenicity within the cupins is confined to the vicilin and legumin seed storage proteins. Profilins are ubiquitous eukaryotic proteins that are nonallergenic, with the exception of profilins from flowering plants. Finally, the Bet v 1 superfamily contains the pathogenesis-related proteins 10 family, the family of major latex proteins and ripening-related proteins, the norcoclaurine synthases, and the cytokinin-binding proteins, with pathogenesis-related proteins 10 family members from certain taxa being the only allergenic members. The study of the distribution of allergenic and nonallergenic members of protein families will provide new insights into the evolution of allergenicity and the factors that make proteins allergenic. (*J Allergy Clin Immunol* ■■■■;■■■:■■■-■■■.)

Key words: Allergens, evolution, protein superfamilies, prolamins, cupins, profilins, Bet v 1

The massive accumulation of sequence data of plant proteins in recent years has made the classification of allergens into protein families possible.^{1,2} It also allowed the identification of a vast range of related proteins that are either not allergenic by nature or that have never come into contact with the human immune system, resulting in a lack of knowledge about their allergenicity. By comparing sequences and structures, related proteins can be grouped together into families, and related families can be grouped together into superfamilies. This much broader picture now allows us to view allergenic proteins in their evolutionary context. The emerging research area of the evolutionary biology of allergens aims to discover

Abbreviations used

ABP: Auxin-binding protein
GLP: Germin-like protein
PR-10: Pathogenesis-related proteins 10

the origin of allergens and thus to elucidate the emergence of allergenicity itself.³

Recently, we have established the AllFam database, a freely available internet resource for classifying allergens into protein families (<http://www.meduniwien.ac.at/allergens/allfam/>).⁴ Querying the database for the distribution of plant food allergens into protein families renders a list that is highly biased toward certain families. As few as 4 protein families contain nearly 60% of all plant food allergens. These families are the prolamin superfamily, which contains, among others, the 2S albumin seed storage proteins and the nonspecific lipid transfer proteins; the cupin superfamily, to which belong the vicilin and legumin-type seed storage globulins; the profilins; and proteins related to the major birch pollen allergen Bet v 1. This review gives a first introduction to the evolutionary scope of the origin of some of the most important plant food allergens.

THE PROLAMIN SUPERFAMILY

Numerically, the prolamin superfamily is the most prominent of all protein families that contain allergenic members. Prolamins are a group of seed storage proteins that are the dominant storage proteins of cereals and other members of the grass family. Exceptions are rice and oats, in which the major storage proteins are 11S globulin like. Prolamins were initially defined on the basis of their solubility in alcohol/water mixtures and their high content of proline and glutamine, hence the name prolamins. Because of their importance in nutrition and food processing, cereal prolamins have been subjected to intense research.⁵ The subgroup of sulfur-rich prolamins consists of a repetitive domain rich in proline and glutamine and a nonrepetitive domain that contains conserved cysteine residues. The identification of the existence of related proteins was based on the presence of this conserved cysteine skeleton.⁶ Three groups of unrelated seed proteins were identified, including the sulfur-rich prolamins and the α -amylase/trypsin inhibitors of cereal seeds, together with the 2S albumins found in a variety of dicotyledonous seeds. This cysteine skeleton has since been

From the Department of Pathophysiology, Medical University of Vienna.

H.B. was supported by the Austrian Science Fund (grant SFB F01802).

Disclosure of potential conflict of interest: The authors have declared that they have no conflict of interest.

Received for publication July 2, 2007; revised July 18, 2007; accepted for publication July 18, 2007.

Reprint requests: Heimo Breiteneder, PhD, Department of Pathophysiology, Medical University of Vienna, AKH-EBO-3Q, Währinger Gürtel 18-20, 1090 Vienna, Austria. E-mail: heimo.breiteneder@meduniwien.ac.at.

0091-6749/\$32.00

© 2007 American Academy of Allergy, Asthma & Immunology

doi:10.1016/j.jaci.2007.07.024

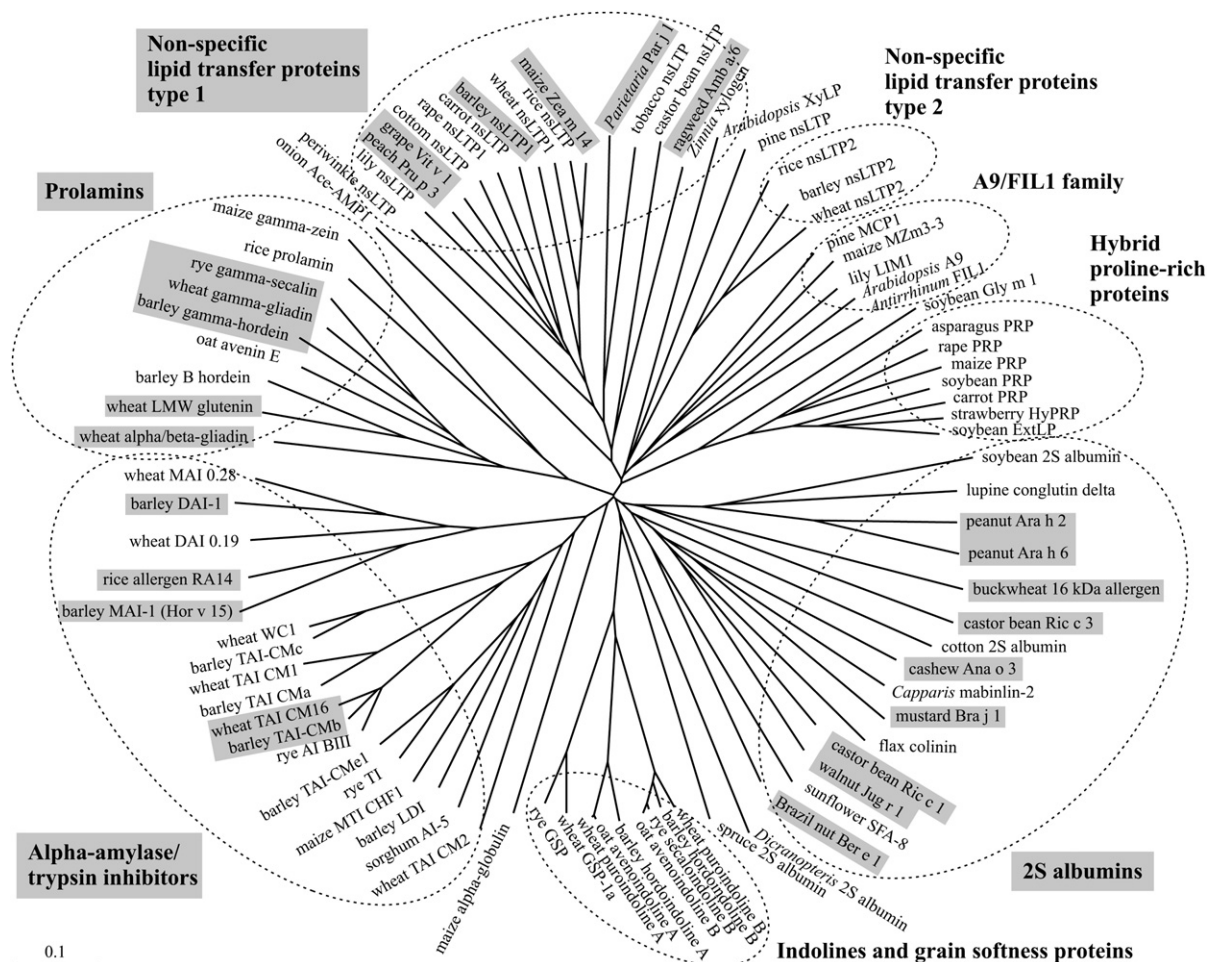


FIG 1. Phylogenetic tree showing the relationships of members of the prolamin superfamily. Allergens and families that contain allergens are shaded. *AI*, α -Amylase inhibitor; *Ace-AMP1*, *Allium cepa* antimicrobial protein 1; *DAI*, dimeric amylase inhibitor; *ExtLP*, extensin-like protein; *GSP*, grain softness protein; *HyPRP*, hybrid proline-rich protein; *LDI*, limit dextrinase inhibitor; *LMW*, low molecular weight; *MAI*, monomeric amylase inhibitor; *MTI*, monomeric trypsin inhibitor; *nsLTP*, nonspecific lipid transfer protein; *PRP*, proline-rich protein; *SFA-8*, sunflower albumin 8; *TAI-CM*, chloroform/methanol soluble tetrameric amylase inhibitor; *TI*, trypsin inhibitor; *WCI*, wheat chymotrypsin inhibitor; *XyLP*, xyloglucan-like protein.

found in a range of small sulfur-rich proteins, most of which are restricted to seeds and some of which are allergens (Fig 1).¹

In addition to the nonspecific lipid transfer proteins that are also members of the prolamin superfamily, indolines represent the other major lipid-binding protein family present in cereal seeds.⁷ Members of the α -globulin family of storage proteins were originally identified in rice⁸ and corn.⁹ The hybrid proline-rich domain B proteins are hydroxyproline-rich cell-wall glycoproteins that have been cloned from different monocotyledonous and dicotyledonous plants. Their C-terminal domain with the 8 cysteines showed similarity to the allergenic hydrophobic protein from soybean.¹⁰ Because the prolamins are essentially defined on the basis of the cysteine skeleton, distantly related members of this superfamily are much more difficult to define than for the cupin superfamily,

which is defined on the basis of related 3-dimensional structures. Thus far, members of the prolamin superfamily have only been reported in seed plants, with the only exception being proteins related to 2S albumins that are present in spores of the ostrich fern.¹¹

Allergenic prolamins include members of the nonspecific lipid transfer protein family,¹² the 2S albumin storage proteins,¹³ the cereal α -amylase/trypsin inhibitors,¹⁴ and the soybean hydrophobic protein.¹⁵ Allergies to cereal prolamin seed storage proteins do not occur very frequently and have been studied mainly in wheat.¹⁶ These protein families have little sequence homology to each other apart from the cysteine skeleton, but they have highly similar α -helical structures. This structure is highly stable to both thermal and proteolytic denaturation, which might contribute to the allergenicity of these proteins.¹⁷

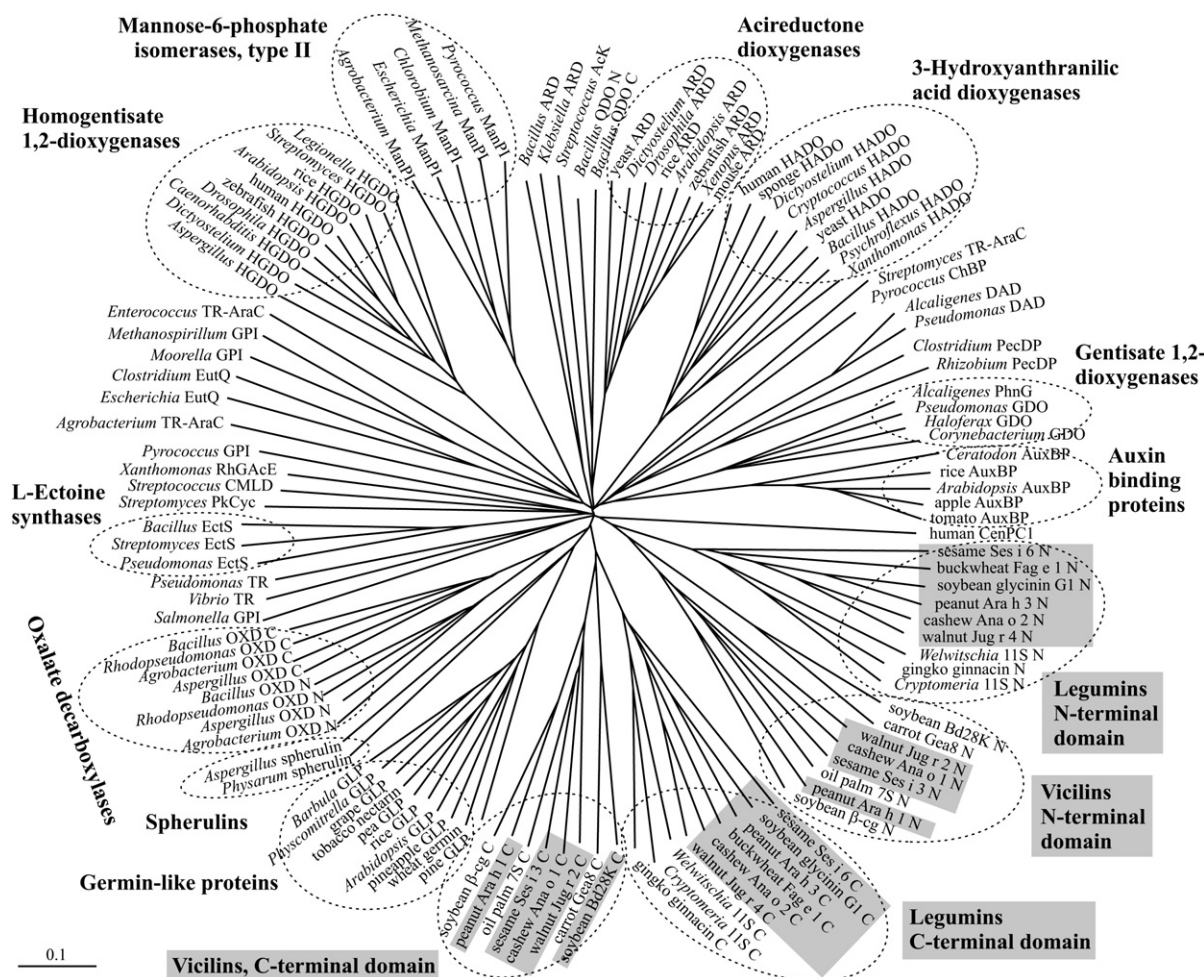


FIG 2. Phylogenetic tree showing the relationships of members of the cupin superfamily. Allergens and families that contain allergens are shaded. *Ack*, Acetate kinase; *ARD*, acireductone dioxygenase; *AuxBP*, auxin-binding protein; β -cg, β -conglycinin; *Bd28K*, 28-kd allergen; *CenPC1*, centromere protein C1; *ChBP*, carbohydrate-binding protein; *CMLD*, 4-carboxymuconolactone decarboxylase; *DAD*, 2,4'-dihydroxyacetophenone dioxygenase; *EctS*, L-ectoine synthase; *EutQ*, ethanolamine utilization protein Q; *GDO*, gentisate 1,2-dioxygenase; *GLP*, germin-like protein; *GPI*, glucose-6-phosphate isomerase; *HADO*, 3-hydroxyanthranilic acid dioxygenase; *HGDO*, homogentisate 1,2-dioxygenase; *ManPI*, mannose-6-phosphate isomerase; *OXD*, oxalate decarboxylase; *PecDP*, pectin degradation protein; *PhnG*, 1-hydroxy-2-naphthoate dioxygenase; *PkCyc*, polyketide cyclase; *RhGAce*, rhamnogalacturonan acetyltransferase; *QDO*, quercetin dioxygenase; *TR*, transcriptional regulator; *TR-AraC*, transcriptional regulator from the AraC family.

THE CUPIN SUPERFAMILY

The cupins are a superfamily of functionally highly diverse proteins whose evolution can be traced from archaea and bacteria to eukaryotes, including plants and animals. Their discovery was based on the identification of sequence similarities between wheat germin, a highly thermostable and protease-resistant glycoprotein; fungal spherulins produced in the slime mold *Physarum polycephalum* on spore formation; and plant globulin seed storage proteins.^{18,19} These proteins were named cupins²⁰ because they share a common architecture described as a 6-stranded β -barrel (Latin “cupa” means barrel). Cupin subgroups are classified on whether these proteins comprise a single cupin domain or whether they have a

duplicated or a multicupin structure. It was estimated that there were a minimum of 18 different functional subclasses based on the proteins’ various enzymatic and non-enzymatic functions (Fig 2).²¹

Monocupins, or single-domain cupins, comprise the majority of cupin proteins and are either monomeric, dimeric, or oligomeric. They mostly comprise enzymes, such as dioxygenases and phosphomannose isomerases.²¹ It is noteworthy that several thermophilic archaea express thermostable cupin domain-containing enzymes, such as the *Pyrococcus furiosus* phosphoglucose isomerase, which shows maximum activity at temperatures in excess of 90°C.²² Auxin-binding proteins (ABPs), a family of dimeric monocupins from plants, are involved in a variety of plant growth responses by interacting with the plant

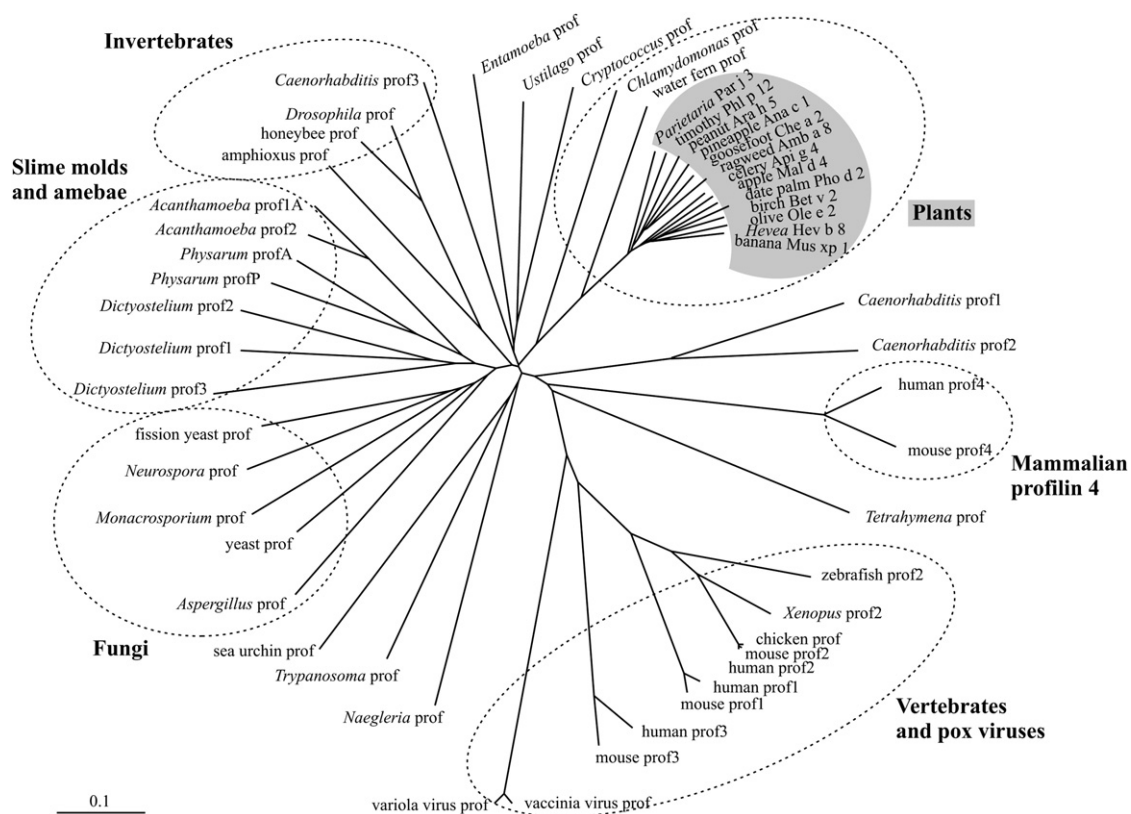


FIG 3. Phylogenetic tree showing the relationships of profilins. Allergens and taxonomic groups that contain allergens are shaded.

hormone auxin.²³ The structure of ABP1 from corn is similar to that of germin, with which it shares 24% sequence identity.²³ ABPs display considerable sequence conservation across a broad range of plants, including apple and strawberry.

Germins and germin-like proteins (GLPs), being oligomeric monocupins, represent the largest family of cupins in plants. Germins from wheat and barley are highly thermostable, hydrogen peroxide-generating oxalate oxidases, with an additional superoxide dismutase activity identified for barley germin.²⁴ The major nectar protein from ornamental tobacco has been identified as a GLP with superoxide dismutase activity that generates high levels of hydrogen peroxide in nectar, thus contributing to defense against potential pathogens.²⁵ Structural analysis revealed the germin protein to be a disc-shaped homohexamer organized as trimers of dimers.²⁴

The first 2-domain cupins, or bicupins, to be discovered were the seed storage proteins of higher plants.¹⁹ They can be classified into trimeric 7S globulins or vicilins and hexameric 11S globulins or legumins. 7S and 11S globulins share relatively low sequence identities of 35% to 45% but display high structural similarity.^{26,27} The discovery of bicupins in prokaryotes corroborates the assumption that these proteins have evolved from the duplication and then fusion of single-domain ancestral prokaryotic genes.²¹ The cupin core provides a stable scaffold that

allows these proteins to survive and function under a great variety of extreme conditions.²⁸ It might well have been the thermotolerance of the cupin tertiary structure that lead to its use by plants as a means for storing amino acids, first in desiccated spores and then in seeds.

The majority of cupin allergens belongs to either the vicilin-like or the legumin-like seed storage protein families and comprise major legume, tree nut, and seed allergens.¹ GLPs have been described as allergens in pepper, orange, and tangerine, apparently owing much of their IgE-binding capacity to their glycosylation.²⁹

THE PROFILINS

Profilins, cytosolic proteins of 12 to 15 kD, are found in all eukaryotic cells and in pox viruses (Fig 3). Profilins bind to monomeric actin and a plethora of other ligands, among them phosphatidyl inositol phosphates and diverse regulatory proteins that contain proline-rich stretches. Profilin is regarded as a key player in the regulation of actin cytoskeleton dynamics during processes such as cell movement, cytokinesis, and signaling.³⁰

Phylogenetic analyses showed that profilins can be divided into 2 groups, one containing vertebrate and viral profilins and the other comprising profilins from invertebrates, protists, plants, and fungi.³¹ Sequence similarity

between profilins from vertebrates and other organisms is low. However, 3-dimensional structures of all profilins are strikingly similar.³² They fold into compact globular mixed α/β structures.

The genomes of most organisms contain small gene families of functionally distinct profilin isoforms. In vertebrates 4 isoforms with highly divergent sequences were found. Isoforms I to III are vertebrate specific and originated early in vertebrate evolution, whereas profilin IV is similar to invertebrate profilins.³³ These profilin isoforms show highly specific tissue expression patterns.³³ Pox virus profilins are similar in sequence to profilin III and are thought to influence the phosphatidyl inositol phosphate metabolism of the host cell.³⁴

In several nonvertebrate organisms, profilin gene families were characterized. Although fungal genomes appear to contain a single isoform, the slime mold *Dictyostelium discoideum* expresses 3 biochemically and functionally distinct profilins.³¹ Organ-specific expression of 3 profilin isoforms was also shown in the nematode *Caenorhabditis elegans*.³⁵ In plants profilin isoforms show low sequence divergence. In the genome of *Arabidopsis*, 8 to 10 profilin genes were found,³⁶ which were classified into 2 groups, one expressed in somatic tissues and the other exclusively expressed late in pollen development.^{36,37} On the basis of multiple sequence alignments, an ancient split between pollen-specific and constitutive isoforms was proposed.³⁶

Allergenic profilins are exclusively found in flowering plants (Fig 3). Because of their great extent of sequence conservation, they constitute a family of highly cross-reactive allergens in monocot and dicot pollens, plant foods, and *Hevea latex*.^{38,39} Interestingly, no allergenic profilins from conifer pollen are known. Likewise, no profilins from sources other than plants have been identified as allergens thus far.

THE BET V 1 SUPERFAMILY

The major birch pollen allergen Bet v 1 is a member of a family of cytoplasmic disease resistance-related proteins.⁴⁰ This family, termed *pathogenesis-related proteins 10* (PR-10),⁴¹ is ubiquitously distributed among dicotyledonous plants⁴² and also present in monocots⁴³ and gymnosperms.⁴⁴

PR-10 proteins are members of small gene families of differentially expressed isoforms. In birch 13 gene loci were detected, 7 of which encoded proteins expressed specifically in pollen.⁴⁵ The existence of stress-induced and developmentally regulated PR-10 isoforms was shown for many different species. The expression of stress-induced PR-10 proteins is triggered by pathogen infection,⁴⁶ plant hormones,⁴³ wounding,⁴⁷ salt stress,⁴⁸ or cold.⁴⁴ Developmentally regulated isoforms that are expressed in high concentrations in pollen, seeds, fruits, or storage organs were identified as allergens cross-reactive with Bet v 1.³⁸ Phylogenetic analyses of PR-10 isoforms revealed that the multiple sequences from each plant family formed a monophyletic group, and sequences from

each family were generally more similar to each other than those from other families, suggesting that the genes of the multigene family underwent strong concerted evolution.⁴²

In recent years a superfamily of Bet v 1-related proteins that extended beyond the PR-10 family was established. A group of proteins from opium poppy specifically expressed in high levels in laticifer cells showed low sequence similarities to PR-10 proteins.⁴⁹ An extended sequence comparison revealed that the Bet v 1 superfamily comprises at least 4 families with low levels of sequence similarity between each other (Fig 4): the PR-10 family, the family of major latex proteins and ripening-related proteins, the norcoclaurine synthases, and cytokinin-binding proteins from legumes.⁵⁰

The major latex protein/ripening-related protein family comprises major latex proteins from laticiferous plants, such as opium poppy and plantain,⁵¹ and proteins expressed in ripening fruits, including wild strawberry⁵² and soybean.⁵³ (S)-norcoclaurine synthases were cloned from meadow rue (*Thalictrum flavum*) and poppy.^{50,54} These enzymes catalyze the first committed step in the biosynthesis of benzyloquinoline alkaloids, such as morphine. Proteins binding the plant hormone cytokinin with sequences related to PR-10 and (S)-norcoclaurine synthases but forming a distinct family were found in the legumes mung bean and lupine.⁵⁵ The identification of a cytokinin-binding PR-10-related protein in the moss *Physcomitrella patens* revealed that the evolutionary origin of the Bet v 1 superfamily can be dated before the development of seed plants.⁵⁶

The first solved 3-dimensional structure of a protein from the Bet v 1 superfamily was that of Bet v 1.⁵⁷ Despite low sequence similarity, the cytokinin-binding protein from mung bean adopts an identical fold.⁵⁵ A striking feature of the Bet v 1 fold is a large cavity that contains 2 tightly bound molecules of deoxycholate in the structure of the Bet v 1 isoform Bet v 1L.⁵⁸ This pointed to a possible biologic function of PR-10 proteins as plant steroid carriers. The PR-10-related protein Hyp1 from St John's wort (*Hypericum perforatum*) catalyzes the final step in the biosynthesis of the antidepressant alkaloid hypericin.⁵⁹ Considering the evolutionary relationship of the PR-10 family to (S)-norcoclaurine synthases and legume cytokinin-binding proteins, it is justified to propose that the Bet v 1 fold evolved as a versatile scaffold that was used by evolution for binding a diverse array of polycyclic, mostly hydrophobic, low-molecular-weight compounds.⁶⁰ In addition, a ribonuclease activity was proposed for several members of the PR-10 family.⁶¹

In recent years, structural comparisons showed that the Bet v 1 fold is also found in diverse protein families outside the plant kingdom. This structural relationship that cannot be detected by means of primary sequence comparison was first shown in 2001, when the structural similarity between PR-10 proteins and the steroidogenic acute regulatory protein-related lipid transfer domain of the human metastatic lymph node protein 64 was discovered.⁶² In the Pfam database of protein families,⁶³ the Bet

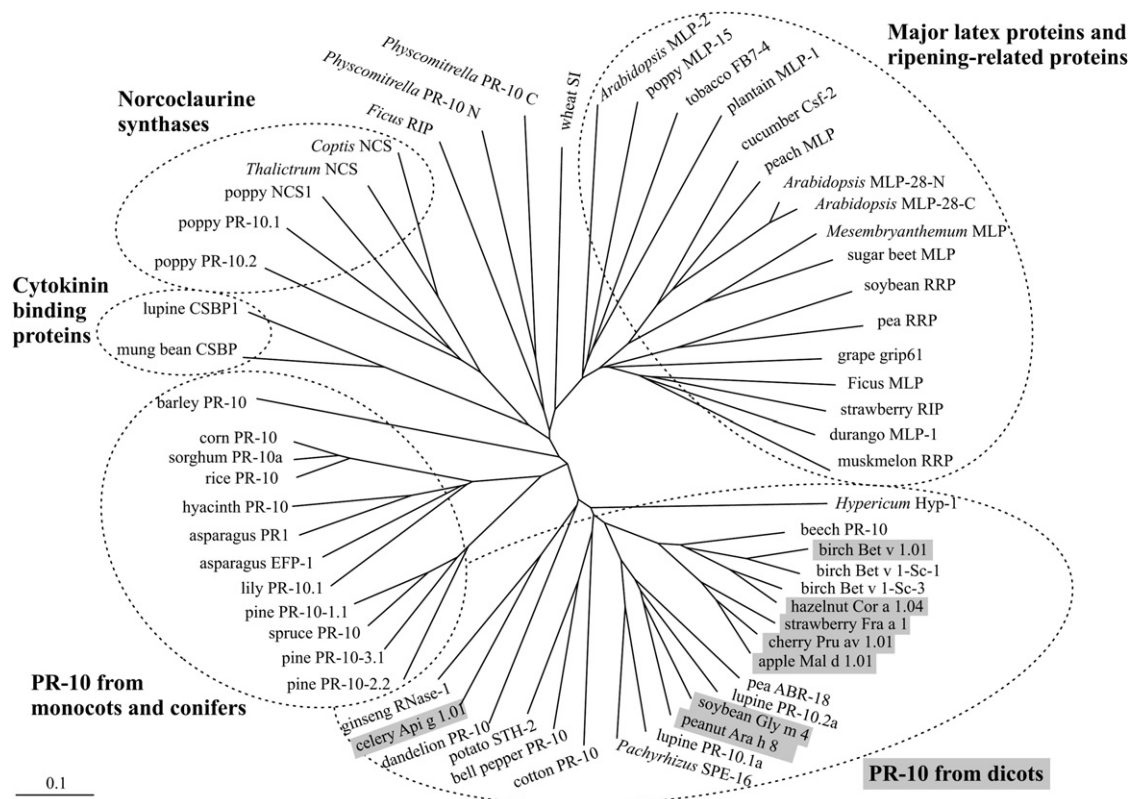


FIG 4. Phylogenetic tree showing the relationships of members of the Bet v 1 superfamily. Allergens and families that contain allergens are shaded. ABR, Absciscic acid-regulated protein; CSBP, cytokinin-specific binding protein; EFP, early flowering protein; FB7, floral bud day 7 explant transcript; MLP, major latex protein; NCS, norcoclaurine synthase; RNase, ribonuclease; RIP, ripening-induced protein; RRP, ripening-related protein; SI, salt-tolerant protein.

v 1 family is classified as member of the “Bet v 1-like” clan, which contains several protein families whose members are found in eukaryotes, bacteria, and archaea and whose functions are in most cases related to lipid binding.

Allergenicity within the Bet v 1 superfamily is confined to the PR-10 family (Fig 4). Cross-reactivity of the major birch pollen allergen Bet v 1 with homologues from plant foods is responsible for birch pollen-associated food allergy.³⁸ The taxonomic distribution of Bet v 1-related allergens is quite limited. Pollen allergens are found exclusively in the order Fagales (birch and beech families), whereas food allergens were described from fruits and vegetables from the families Rosaceae (apple and stone fruits), Apiaceae (celery and carrot), and Fabaceae (peanut and soybean).

CONCLUSIONS

There is an ongoing discussion about which features make a protein allergenic. One view claims that any immunogenic protein presented in the right context and in sufficient quantity to the immune system of an atopic individual can become an allergen.⁶⁴ The other view attributes allergenicity to certain structural features of the

proteins.^{1,17} For this discussion, the distribution of allergenic and nonallergenic members in the prolamin and Bet v 1 superfamilies could provide further insights. In cereal grains certain prolamins and bifunctional inhibitors are allergenic, whereas indolines and the cereal α -globulins are not, although they are expressed in high amounts. Likewise, fruits and legume seeds contain allergenic PR-10 proteins, as well as nonallergenic ripening-related proteins, both members of the Bet v 1 superfamily.

The study of the evolutionary biology of allergens will eventually lead to the study of the evolution of allergenicity. Two main scenarios are imaginable. In scenario 1 allergenicity has been an intrinsic molecular property of the ancestral members of certain protein families and has been distributed into the present-day allergen sources of the plant kingdom through evolutionary processes. Scenario 2 would be that allergenicity emerged much later, again as structural features but randomly evolving only in certain proteins and inherited by their progeny. Both scenarios are at present highly speculative. The distribution of allergens in the prolamin superfamily (Fig 1) supports scenario 1. Allergenic proteins can be found in many of the member families of this superfamily, indicating a tendency of a protoprolamin to act as an allergen. In contrast, the other superfamilies discussed in this review

show a very limited distribution of allergens among their member families (Figs 2-4), which would be more in accordance with scenario 2. For example, profilins from invertebrates, fungi, and plants show very low sequence similarities to human profilins; however, only plant profilins have been described as allergens (Fig 3).

In conclusion, the introduction of evolutionary concepts into allergen research places allergenic proteins in a much larger context that will stimulate further experiments that might eventually lead to a new understanding of allergenicity.

REFERENCES

- Breiteneder H, Radauer C. A classification of plant food allergens. *J Allergy Clin Immunol* 2004;113:821-30.
- Radauer C, Breiteneder H. Pollen allergens are restricted to few protein families and show distinct patterns of species distribution. *J Allergy Clin Immunol* 2006;117:141-7.
- Mills EN, Jenkins JA, Alcocer MJ, Shewry PR. Structural, biological, and evolutionary relationships of plant food allergens sensitizing via the gastrointestinal tract. *Crit Rev Food Sci Nutr* 2004;44:379-407.
- Radauer C, Marl A, Breiteneder H. The AllFam database—a resource for classifying allergenic proteins. *Allergy* 2007;62(suppl 83):154.
- Shewry PR, Napier JA, Tatham AS. Seed storage proteins: structures and biosynthesis. *Plant Cell* 1995;7:945-56.
- Kreis M, Forde BG, Rahman S, Mifflin BJ, Shewry PR. Molecular evolution of the seed storage proteins of barley, rye and wheat. *J Mol Biol* 1985;183:499-502.
- Douliez A-P, Michon T, Elmorjani K, Marion D. Structure, biological and technological functions of lipid transfer proteins and indolines, the major lipid binding proteins from cereal kernels. *J Cereal Sci* 2000;32:1-20.
- Nakase M, Hotta H, Adachi T, Aoki N, Nakamura R, Masumura T, et al. Cloning of the rice seed alpha-globulin-encoding gene: sequence similarity of the 5'-flanking region to those of the genes encoding wheat high-molecular-weight glutenin and barley D hordein. *Gene* 1996;170:223-6.
- Woo YM, Hu DW, Larkins BA, Jung R. Genomics analysis of genes expressed in maize endosperm identifies novel seed proteins and clarifies patterns of zein gene expression. *Plant Cell* 2001;13:2297-317.
- José-Estanyol M, Puigdomenech P. Plant cell wall glycoproteins and their genes. *Plant Physiol Biochem* 2000;38:97-108.
- Rodin J, Rask L. Characterization of matheuccin, the 2.2S storage protein of the ostrich fern. Evolutionary relationship to angiosperm seed storage proteins. *Eur J Biochem* 1990;192:101-7.
- Breiteneder H, Mills C. Nonspecific lipid-transfer proteins in plant foods and pollens: an important allergen class. *Curr Opin Allergy Clin Immunol* 2005;5:275-9.
- Pastorello EA, Pompei C, Pravettoni V, Brenna O, Farioli L, Trambaioli C, et al. Lipid transfer proteins and 2S albumins as allergens. *Allergy* 2001;56(suppl 67):45-7.
- Pastorello EA, Farioli L, Conti A, Pravettoni V, Bonomi S, Iametti S, et al. Wheat IgE-mediated food allergy in European patients: alpha-amylase inhibitors, lipid transfer proteins and low-molecular-weight glutenins. Allergenic molecules recognized by double-blind, placebo-controlled food challenge. *Int Arch Allergy Immunol* 2007;144:10-22.
- Baud F, Pebay-Peyroula E, Cohen-Addad C, Odani S, Lehmann MS. Crystal structure of hydrophobic protein from soybean; a member of a new cysteine-rich family. *J Mol Biol* 1993;231:877-87.
- Maruyama N, Ichise K, Katsube T, Kishimoto T, Kawase S, Matsumura Y, et al. Identification of major wheat allergens by means of the *Escherichia coli* expression system. *Eur J Biochem* 1998;255:739-45.
- Breiteneder H, Mills EN. Molecular properties of food allergens. *J Allergy Clin Immunol* 2005;115:14-23.
- Lane BG, Bernier F, Dratewka-Kos E, Shafai R, Kennedy TD, Pyne C, et al. Homologies between members of the germin gene family in hexaploid wheat and similarities between these wheat germ proteins and certain *Physarum sphaerulins*. *J Biol Chem* 1991;266:10461-9.
- Bäumlein H, Braun H, Kakhovskaya IA, Shutov AD. Seed storage proteins of spermatophytes share a common ancestor with desiccation proteins of fungi. *J Mol Evol* 1995;41:1070-5.
- Dunwell JM. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. *Biotechnol Genet Eng Rev* 1998;15:1-32.
- Dunwell JM, Purvis A, Khuri S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* 2004;65:7-17.
- Verhees CH, Huynen MA, Ward DE, Schiltz E, de Vos WM, van der Oost J. The phosphoglucose isomerase from the hyperthermophilic archaeon *Pyrococcus furiosus* is a unique glycolytic enzyme that belongs to the cupin superfamily. *J Biol Chem* 2001;276:40926-32.
- Woo EJ, Marshall J, Baulby J, Chen JG, Venis M, Napier RM, et al. Crystal structure of auxin-binding protein 1 in complex with auxin. *EMBO J* 2002;21:2877-85.
- Woo EJ, Dunwell JM, Goodenough PW, Marvier AC, Pickersgill RW. Germin is a manganese containing homohexamer with oxalate oxidase and superoxide dismutase activities. *Nat Struct Biol* 2000;7:1036-40.
- Thornburg RW, Carter CE, Powell A, Mittler R, Rizhsky L, Horner HT. A major function of the tobacco floral nectary is defense against microbial attack. *Plant System Evolution* 2003;238:211-8.
- Maruyama N, Adachi M, Takahashi K, Yagasaki K, Kohno M, Takenaka Y, et al. Crystal structures of recombinant and native soybean beta-conglycinin beta homotrimers. *Eur J Biochem* 2001;268:3595-604.
- Adachi M, Kanamori J, Masuda T, Yagasaki K, Kitamura K, Mikami B, et al. Crystal structure of soybean 11S globulin: glycinin A3B4 homohexamer. *Proc Natl Acad Sci U S A* 2003;100:7395-400.
- Thompson MJ, Eisenberg D. Transproteomic evidence of a loop-deletion mechanism for enhancing protein thermostability. *J Mol Biol* 1999;290:595-604.
- Polt G, Ahrazem O, Paschinger K, Ibanez MD, Salcedo G, Wilson IB. Molecular and immunological characterization of the glycosylated orange allergen Cit s 1. *Glycobiology* 2007;17:220-30.
- Witke W. The role of profilin complexes in cell motility and other cellular processes. *Trends Cell Biol* 2004;14:461-9.
- Arasada R, Gloss A, Tunggal B, Joseph JM, Rieger D, Mondal S, et al. Profilin isoforms in *Dictyostelium discoideum*. *Biochim Biophys Acta* 2007;1773:631-41.
- Thorn KS, Christensen HE, Shigeta R, Huddler D, Shalaby L, Lindberg U, et al. The crystal structure of a major allergen from plants. *Structure* 1997;5:19-32.
- Polet D, Lambrechts A, Vandepoele K, Vandekerckhove J, Ampe C. On the origin and evolution of vertebrate and viral profilins. *FEBS Lett* 2007;581:211-7.
- Machesky LM, Cole NB, Moss B, Pollard TD. Vaccinia virus expresses a novel profilin with a higher affinity for polyphosphoinositides than actin. *Biochemistry* 1994;33:10815-24.
- Polet D, Lambrechts A, Ono K, Mah A, Peelman F, Vandekerckhove J, et al. *Caenorhabditis elegans* expresses three functional profilins in a tissue-specific manner. *Cell Motil Cytoskeleton* 2006;63:14-28.
- Huang S, McDowell JM, Weise MJ, Meagher RB. The *Arabidopsis* profilin gene family. Evidence for an ancient split between constitutive and pollen-specific profilin genes. *Plant Physiol* 1996;111:115-26.
- Kandasamy MK, McKinney EC, Meagher RB. Plant profilin isoforms are distinctly regulated in vegetative and reproductive tissues. *Cell Motil Cytoskeleton* 2002;52:22-32.
- 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.
- Radauer C, Willeroider M, Fuchs H, Hoffmann-Sommergruber K, Thalhammer J, Ferreira F, et al. Cross-reactive and species-specific immunoglobulin E epitopes of plant profilins: an experimental and structure-based analysis. *Clin Exp Allergy* 2006;36:920-9.
- Breiteneder H, Pottenburger K, Bito A, Valenta R, Kraft D, Rumpold H, et al. The gene coding for the major birch pollen allergen Betv1, is highly homologous to a pea disease resistance response gene. *EMBO J* 1989;8:1935-8.
- van Loon LC, Rep M, Pieterse CM. Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* 2006;44:135-62.
- Wen J, Vanek-Krebitz M, Hoffmann-Sommergruber K, Scheiner O, Breiteneder H. The potential of Betv1 homologues, a nuclear multigene

- family, as phylogenetic markers in flowering plants. *Mol Phylogenet Evolution* 1997;8:317-33.
43. Wang CS, Huang JC, Hu JH. Characterization of two subclasses of PR-10 transcripts in lily anthers and induction of their genes through separate signal transduction pathways. *Plant Mol Biol* 1999;40:807-14.
 44. Yu XS, Ekramoddoullah AKM, Misra S. Characterization of Pin m III cDNA in western white pine. *Tree Physiol* 2000;20:663-71.
 45. Schenk MF, Gilissen LJ, Esselink GD, Smulders MJ. Seven different genes encode a diverse mixture of isoforms of Bet v 1, the major birch pollen allergen. *BMC Genomics* 2006;7:168-82.
 46. Swoboda I, Scheiner O, Heberle-Bors E, Vicente O. cDNA cloning and characterization of three genes in the Bet v 1 gene family that encode pathogenesis-related proteins. *Plant Cell Environ* 1995;18:865-74.
 47. Liu JJ, Ekramoddoullah AK, Piggott N, Zamani A. Molecular cloning of a pathogen/wound-inducible PR10 promoter from *Pinus monticola* and characterization in transgenic *Arabidopsis* plants. *Planta* 2005;221:159-69.
 48. Jain S, Srivastava S, Sarin NB, Kav NN. Proteomics reveals elevated levels of PR 10 proteins in saline-tolerant peanut (*Arachis hypogaea*) calli. *Plant Physiol Biochem* 2006;44:253-9.
 49. Osmark P, Boyle B, Brisson N. Sequential and structural homology between intracellular pathogenesis-related proteins and a group of latex proteins. *Plant Mol Biol* 1998;38:1243-6.
 50. Liscombe DK, MacLeod BP, Loukanina N, Nandi OI, Facchini PJ. Evidence for the monophyletic evolution of benzyloquinoline alkaloid biosynthesis in angiosperms. *Phytochemistry* 2005;66:2501-20.
 51. Nessler CL, Allen RD, Galewsky S. Identification and Characterization of Latex-Specific proteins in opium poppy. *Plant Physiol* 1985;79:499-504.
 52. Nam YW, Tichit L, Leperlier M, Cuerq B, Marty I, Lelievre JM. Isolation and characterization of mRNAs differentially expressed during ripening of wild strawberry (*Fragaria vesca* L.) fruits. *Plant Mol Biol* 1999;39:629-36.
 53. Stromvik MV, Sundararaman VP, Vodkin LO. A novel promoter from soybean that is active in a complex developmental pattern with and without its proximal 650 base pairs. *Plant Mol Biol* 1999;41:217-31.
 54. Samanani N, Liscombe DK, Facchini PJ. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzyloquinoline alkaloid biosynthesis. *Plant J* 2004;40:302-13.
 55. Pasternak O, Bujacz GD, Fujimoto Y, Hashimoto Y, Jelen F, Otlewski J, et al. Crystal structure of *Vigna radiata* cytokinin-specific binding protein in complex with zeatin. *Plant Cell* 2006;18:2622-34.
 56. Gonneau M, Pagant S, Brun F, Laloue M. Photoaffinity labelling with the cytokinin agonist azido-CPPU of a 34 kDa peptide of the intracellular pathogenesis-related protein family in the moss *Physcomitrella patens*. *Plant Mol Biol* 2001;46:539-48.
 57. Gajhede M, Osmark P, Poulsen FM, Ipsen H, Larsen JN, van Neerven RJJ, et al. X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nat Struct Biol* 1996;3:1040-5.
 58. Markovic-Housley Z, Degano M, Lamba D, von Roepenack-Lahaye E, Clemens S, Susani M, et al. Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier. *J Mol Biol* 2003;325:123-33.
 59. Bais HP, Vepachedu R, Lawrence CB, Stermitz FR, Vivanco JM. Molecular and biochemical characterization of an enzyme responsible for the formation of hypericin in St. John's wort (*Hypericum perforatum* L.). *J Biol Chem* 2003;278:32413-22.
 60. Koistinen KM, Soininen P, Venalainen TA, Hayrinen J, Laatikainen R, Perakyla M, et al. Birch PR-10c interacts with several biologically important ligands. *Phytochemistry* 2005;66:2524-33.
 61. Park CJ, Kim KJ, Shin R, Park JM, Shin YC, Paek KH. Pathogenesis-related protein 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway. *Plant J* 2004;37:186-98.
 62. Neudecker P, Schweimer K, Nerkamp J, Scheurer S, Vieths S, Sticht H, et al. Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1. *J Biol Chem* 2001;276:22756-63.
 63. Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, et al. Pfam: clans, web tools and services. *Nucleic Acids Res* 2006;34:D247-51.
 64. Aalberse RC. Structural features of allergenic molecules. *Chem Immunol Allergy* 2006;91:134-46.