

2015

An Evolutionary Perspective on Human Cross-sensitivity to Tree Nut and Seed Allergens

Amanda E. Fisher

Rancho Santa Ana Botanic Garden, Claremont, California

Annalise M. Nawrocki

Pomona College, Claremont, California

Follow this and additional works at: <https://scholarship.claremont.edu/aliso>



Part of the [Botany Commons](#), [Evolution Commons](#), and the [Nutrition Commons](#)

Recommended Citation

Fisher, Amanda E. and Nawrocki, Annalise M. (2015) "An Evolutionary Perspective on Human Cross-sensitivity to Tree Nut and Seed Allergens," *Aliso: A Journal of Systematic and Floristic Botany*. Vol. 33: Iss. 2, Article 3.

Available at: <https://scholarship.claremont.edu/aliso/vol33/iss2/3>

AN EVOLUTIONARY PERSPECTIVE ON HUMAN CROSS-SENSITIVITY TO TREE NUT AND SEED ALLERGENS

AMANDA E. FISHER¹⁻³ AND ANNALISE M. NAWROCKI²

¹*Rancho Santa Ana Botanic Garden and Claremont Graduate University, 1500 North College Avenue, Claremont, California 91711 (Current affiliation: Department of Biological Sciences, California State University, Long Beach, 1250 Bellflower Boulevard, Long Beach, California 90840);* ²*Pomona College, 333 North College Way, Claremont, California 91711 (Current affiliation: Amgen Inc., annarocky@gmail.com)*

³*Corresponding author (Amanda.Fisher@csulb.edu)*

ABSTRACT

Tree nut allergies are some of the most common and serious allergies in the United States. Patients who are sensitive to nuts or to seeds commonly called nuts are advised to avoid consuming a variety of different species, even though these may be distantly related in terms of their evolutionary history. This is because studies in the literature report that patients often display sensitivity to multiple nut species (cross-sensitivity) if they have an existing nut allergy. These reports suggest that cross-sensitivity in patients with nut allergies may be caused by an IgE antibody reacting with epitopes present in the seed proteins of different species (cross-reactivity), for example, if IgE isolated from the serum of a patient were able to bind to both almond and peanut allergens. We hypothesize that allergenic proteins in seeds may have similar amino acid sequences that cause the observed cross-sensitivity. Here, we test the hypothesis that similarity in the protein sequences of allergenic nuts drives cross-sensitivity and cross-reactivity by reconstructing the gene trees of three allergenic seed-storage proteins (vicilin, legumin, and 2S albumin) from species sampled across vascular plants. We generate estimates of their phylogenetic relationships and compare these to the allergen cross-sensitivity and cross-reactivity data that is reported in the literature. In general, evolutionary relationships of the three proteins are congruent with the current understanding of plant species relationships. However, we find little evidence that distantly related nut species reported to be cross-reactive share similar vicilin, legumin, or 2S albumin amino acid sequences. Our data thus suggests that features of the proteins other than their amino acid sequences may be driving the cross-reactivity observed during *in vitro* tests and skin tests. Our results support current treatment guidelines to limit nut and seed consumption if allergies are present in a patient. More studies are necessary to better understand the characteristics of allergenic proteins and patterns of cross-sensitivity in patients who suffer from nut allergies.

Key words: 2S albumin, cupins, evolutionary medicine, legumin, nut allergy, seed storage proteins, vicilin.

INTRODUCTION

Nuts are a major agricultural commodity in California (USDA 2014) and nut consumption has been shown to lower cholesterol (Morgan and Clayshulte 2000; Garg et al. 2003), reduce the incidence of coronary heart disease (Fraser et al. 1992), and lessen the impact of age-related brain dysfunction (Carey et al. 2012). Despite their potential health benefits, nut allergies are among the most common allergies in the United States (Bock et al. 2001), and consuming nuts may elicit serious and life-threatening immunological responses in people with food allergies (Teuber et al. 2003; Cianferoni and Muraro 2012).

Food allergies are a major topic of research in immunology and there are a number of immunological terms used throughout the paper that we have defined in a glossary (Appendix 1). Research on food allergies is complicated by the observation that proteins that have been identified as allergens

in some species are not necessarily allergens in all species in which they are found. For example, vicilin protein has been identified as an allergen in peanut and it is also found in kiwi seeds, but patients with peanut allergies do not typically also have a kiwi allergy. However, a patient who exhibits an allergy to a protein in one plant species may also demonstrate cross-sensitivity to another species (Appendix 1), presumably due to the similarity of proteins the two plants contain. For example, patients with tree nut allergies may also show cross-sensitivity to peanuts (Ewan 1996; Teuber et al. 2003). Patients who exhibit sensitivity to one species of nut are thus advised to avoid all tree nuts and peanuts (Ewan 1996); but, it is unknown if this level of caution is warranted, as nuts and edible seeds have evolved multiple times in the plant tree of life (Fig. 1).

Nut is a botanical term that describes a fruit containing a single seed, with a hard, dry, outer layer, and a special covering called a cupule (the "cap" of an acorn) (Brouk 1975; Harris and Harris 2001). Fruit type is not generally conserved within plant lineages, but true nuts are only found in the plant order Fagales, which contains commonly consumed tree nuts such as pecans, walnuts, and hazelnuts (Fig. 1). Many fruits colloquially referred to as nuts are actually other types of fruits (e.g., coconuts) or seeds that have a hard covering (e.g.,

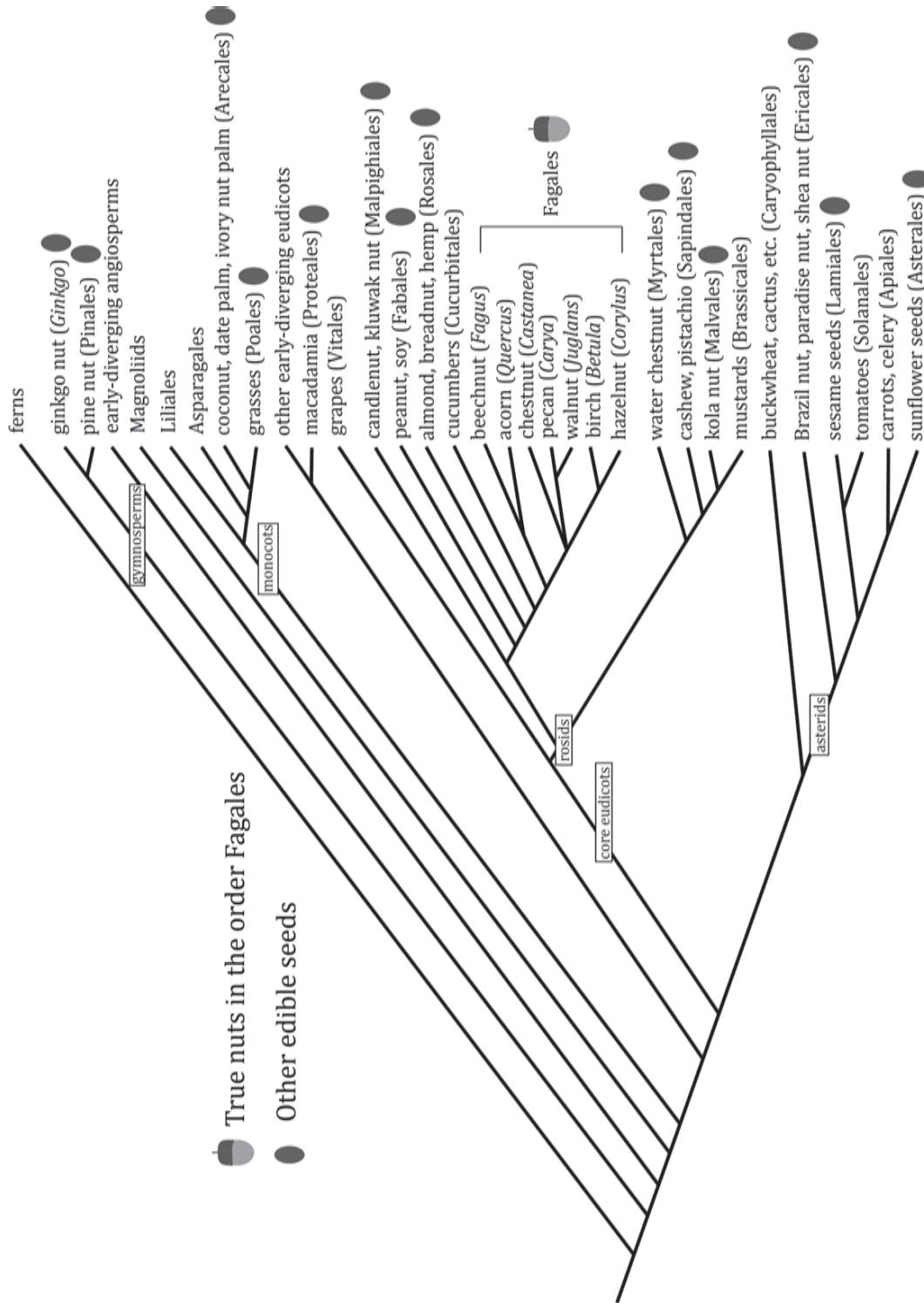


Fig. 1. Evolutionary relationships of nuts and edible seeds on a summary tree of vascular plant evolution. Not all plant orders are shown; notably missing are most of the early-diverging eudicots and core eudicots. The true nuts are restricted to the Fagales order (acorns, chestnuts, hazelnuts, pecans, walnuts) and are indicated with an acorn. Plant orders with seeds that are colloquially called nuts are labeled with a gray oval. For example, peanuts are considered nuts (and one of their common names is "ground nuts"), but botanically-speaking they are the seed of a legume. Adapted from the Angiosperm Phylogeny Website and others (Li et al. 2004; Stevens 2001 onwards).

Table 1. Common nuts and edible seeds according to their classification in the plant tree of life. Fruit structure is not highly conserved in major plant lineages, but true nuts are only found in the Fagales. The table is not a comprehensive list of nuts and edible seeds and it is decidedly biased towards those eaten in the US and Europe. An asterisk next to the common name indicates there is a published genome or other genetic resource, such as a transcriptome, for this species.

Common name	Order	Family	Species	Fruit type
Coconut	Arecales	Arecaceae	<i>Cocos nucifera</i> L.	Drupe
Date palm nut*	Arecales	Arecaceae	<i>Phoenix dactylifera</i> L.	Drupe
Ivory nut palm	Arecales	Arecaceae	<i>Phytelephas aequatorialis</i> Spruce	Dry berry
Sunflower seed*	Asterales	Asteraceae	<i>Helianthus annuus</i> L.	Achene
Brazil nut	Ericales	Lecythidaceae	<i>Bertholletia excelsa</i> Bonpl.	Capsule
Paradise nut	Ericales	Lecythidaceae	<i>Lecythis zabucajo</i> Aubl.	Capsule
Shea nut	Ericales	Sapotaceae	<i>Vitellaria paradoxa</i> C.F. Gaertn.	Drupe
Tahitian chestnut, mape	Fabales	Fabaceae	<i>Inocarpus fagifer</i> (Parkinson ex Zollinger) Fosberg	Indehiscent, 1-seeded legume?
Peanut*	Fabales	Fabaceae	<i>Arachis hypogaea</i> L.	Legume
Soybean*	Fabales	Fabaceae	<i>Glycine max</i> (L.) Merr.	Legume
Acorn*	Fagales	Fagaceae	<i>Quercus</i> sp.	Nut
Beechnut*	Fagales	Fagaceae	<i>Fagus</i> sp.	Nut
Butternut	Fagales	Juglandaceae	<i>Juglans cinerea</i> L.	Nut
Chestnut	Fagales	Fagaceae	<i>Castanea sativa</i> Mill.	Nut
Hazelnut, filbert	Fagales	Betulaceae	<i>Corylus avellana</i> L., <i>Corylus americana</i> Walter	Nut
Hickory nut	Fagales	Juglandaceae	<i>Carya</i> sp. Nutt.	Tryma
Pecan	Fagales	Juglandaceae	<i>Carya illinoensis</i> (Wangenh.) K. Koch	Tryma, drupaceous nut
English/Persian walnut	Fagales	Juglandaceae	<i>Juglans regia</i> L.	Tryma, drupaceous nut
Black walnut	Fagales	Juglandaceae	<i>Juglans nigra</i> L.	Tryma, drupaceous nut
California black walnut	Fagales	Juglandaceae	<i>Juglans californica</i> S. Watson	Tryma, drupaceous nut
Ginkgo nut*	Ginkgoales	Ginkgoaceae	<i>Ginkgo biloba</i> L.	Seed with a fleshy covering
Candlenut, kukui nut	Malpighiales	Euphorbiaceae	<i>Aleurites moluccana</i> (L.) Willd.	Drupe
Kluwak nut	Malpighiales	Salicaceae	<i>Pangium edule</i> Reinw.	Berry
Kola nut	Malvales	Malvaceae	<i>Cola acuminata</i> Schott. & Endl.	Follicle
Water chestnut	Myrtales	Lythraceae	<i>Trapa natans</i> L.	Pseudo-drupe
Pine nut*, piñon, pignolia, Italian stone pine nut, Chinese pine nut, Mexican piñon nut	Pinales	Pinaceae	<i>Pinus edulis</i> Engelm., <i>Pinus pinea</i> L., <i>Pinus koraiensis</i> Siebold & Zucc., <i>P. cembroides</i> Zucc.	Seeds develop in a woody cone
Macadamia	Proteales	Proteaceae	<i>Macadamia integrifolia</i> Maiden & Betche, <i>Macadamia tetraphylla</i> L.A.S. Johnson	Follicle
Almond*	Rosales	Rosaceae	<i>Prunus dulcis</i> (Mill.) D.A. Webb	Drupe
Breadnut	Rosales	Moraceae	<i>Brosimum alicastrum</i> Sw., <i>Artocarpus camansi</i> Blanco	Drupe, sorosis
Hemp nut*	Rosales	Cannabaceae	<i>Cannabis sativa</i> L.	Achene
Cashew	Sapindales	Anacardiaceae	<i>Anacardium occidentale</i> L.	Dry drupe or drupaceous nut
Lychee nut*	Sapindales	Sapindaceae	<i>Litchi chinensis</i> Sonn.	Drupe
Pilnut	Sapindales	Burseraceae	<i>Canarium ovatum</i> Engl.	Drupe
Pistachio	Sapindales	Anacardiaceae	<i>Pistacia vera</i> L.	Drupe

almonds, Brazil nuts, cashews, macadamia nuts, peanuts, pine nuts, pistachios). We use the term nut in the colloquial sense throughout this paper. Even when we consume a botanically true nut, we are only ingesting the seed contained within the nut. Edible seeds develop within many different types of fruits (Table 1) and may be only distantly related to true nut-bearing plants found in the order Fagales (Fig. 1). Edible seeds are found in all of the major lineages of seed plants, including gymnosperms (pine nuts), monocots (e.g., coconuts and grasses), and eudicots (e.g., almonds, Brazil nuts, peanuts). These three plant lineages last shared a recent common ancestor at least 300 million years ago (Stein et al. 2012; Magallon et al. 2013), approximately the same amount of time since the common ancestor of amphibians and mammals

(Hedges 2009). The evolutionary distance between species of plants that produce edible seeds may decrease the likelihood that their seeds contain similar proteins, or that proteins shared between these species have retained sequence similarity.

The nuts most commonly consumed in the United States and Europe are almonds, cashews, hazelnuts, peanuts, pecans, and walnuts. Less frequently eaten are Brazil nuts, chestnuts, macadamia nuts, pine nuts, pistachios, and others (Table 1). These nuts contain thousands of different proteins (Clarke et al. 2000) and several of the proteins are potential allergens that are either plant defense proteins to protect against fungi, bacteria, viruses, and invertebrates, or seed storage proteins that provide nutrition for the germinating embryo (Table 2; Radauer and Breiteneder 2007). Proteins are chains of amino

Table 2. Proteins that have been identified as allergens in edible nuts and seeds.

NUT AND SEED ALLERGENS							
	Vicilin (7S)	Legumin (11S)	2S albumin	Lipid- transfer- protein (LTP)	Profilin	Other	References
COMMON ALLERGENS							
Almond (<i>Prunus dulcis</i>)		Y	Y	Y	Y		(Allergen.org 2014); legumin: (Willison et al. 2011); 2S albumin: (Poltronieri et al. 2002)
Brazil nut (<i>Bertholletia excelsa</i>)		Y	Y				(Altenbach et al. 1987); (Roux et al. 2003); legumin: Beyer et al. unpubl. ¹ ; 2S albumin: (Pastorello et al. 2001)
Cashew (<i>Anacardium occidentale</i>)	Y	Y	Y				Vicilin and legumin: (Wang et al. 2002); 2S albumin: (Robotham et al. 2005)
English walnut (<i>Juglans regia</i>)	Y	Y	Y	Y			(Roux et al. 2003); vicilin & 2S albumin: (Teuber et al. 1998); LTP: (Pastorello et al. 2001); (Teuber et al. 1999; Teuber et al. 2003); legumin: (Wallowitz et al. 2006)
Hazelnut/filbert (<i>Corylus avellana</i>)	Y	Y	Y	Y	Y	Bet v 1-like, oleosin	(Roux et al. 2003; Allergen.org 2014); legumin: (Beyer et al. 2002); LTP & Bet v 1-like: (Pastorello et al. 2001)
Peanut (<i>Arachis hypogaea</i>)	Y	Y	Y	Y	Y	Bet v 1-like, oleosin, defensin	(Allergen.org 2014); vicilin: (Burks et al. 1995); legumin: (Rabjohn et al. 1999); 2S albumin & profilin: (Kleber-Janke et al. 1999); LTP: (Krause et al. 2009); Bet v 1: (Mittag et al. 2004)
Soybean (<i>Glycine max</i>)	Y	Y	Y		Y	Defensin, Bet v 1-like	(Allergen.org 2014); 2S albumin: (Ebisawa et al. 2013); profilin: (Rihs et al. 1999); defensin: (Codina et al. 1997 <i>a, b</i>); Bet v 1-like: (Crowell et al. 1992)
Wheat (<i>Triticum aestivum</i>)				Y	Y	ω 5-gliadin, glutenin	LTP: (Sander et al. 2011); profilin: (Rihs et al. 1994); ω 5-gliadin: (Lehto et al. 2010); glutenin: (Anderson et al. 1999)
LESS COMMON ALLERGENS							
Black walnut (<i>Juglans nigra</i>)	Y		Y				(Yunginger 1988; Bannon et al. 2001); vicilin & 2S albumin: Ling et al. unpubl. ²
Chestnut (<i>Castanea sativa</i>)				Y			(Roux et al. 2003); LTP: (Diaz-Perales et al. 2000)
Pistachio (<i>Pistacia vera</i>)	Y	Y	Y			Manganese superoxide dismutase	(Allergen.org 2014); vicilin: (Willison et al. 2008)

¹ BEYER, K., L. BARDINA, G. GRISHINA, A. ASHRAF, S. S. TEUBER, B. NIGGEMANN, AND H. A. SAMPSON. Identification of an 11S globulin as a Brazil nut food allergen. www.allergen.org.² LING, M., J. YE, G. COCKRELL, H. SAMPSON, G. A. BANNON, J. S. STANLEY, AND A.W. BURKS. Characterization of two major allergens in English and black walnut. www.allergen.org.

Table 3. Model selection for gene tree analyses. Model reported and used here represents the best-fit model for each protein that was available as a model choice in RAxML.

Protein	Model	DeltaAIC	AIC	-lnL
Legumin	LG + I + Γ	206.02	47015.20	-23376.60
Vicilin	LG + I + Γ	24.18	31712.00	-15757.00
2S albumin	JTT + I + Γ	700.55	14764.12	-7258.06

acids that fold into three-dimensional structures. Antibodies in the human immune system can interact with short sections of the amino acid sequence called epitopes. A majority of the proteins in seeds can be grouped into four protein families: prolamins, cupins, profilins, and the Bet v 1-like family. The latter is a protein family named after the birch pollen allergen recovered from the white birch, *Betula verrucosa* Ehrh. Members of these protein families have different structural features that are described below (and reviewed in Breiteneder and Radauer 2004), but they are all highly resistant to water stress and to thermal and proteolytic denaturation, and these characteristics may contribute to their allergenicity.

Members of several protein families have been implicated in allergic reactions to nuts; these include prolamins, seed storage proteins in the cupin superfamily (including legumins [11S globulins] and vicilins [7S or 8S globulins]), and profilins (Witke 2004). Prolamins are the major protein found in grains such as wheat and corn, and contain the lipid-transfer and 2S albumin proteins. The amino acid sequences of these proteins are highly variable, but they share a pattern of 6–8 conserved cysteine residues and a conserved three-dimensional structure (Mills et al. 2004). The prolamin family is only found in the land plants (flowering plants, gymnosperms, ferns, and mosses) and most plant proteins that are allergenic when ingested are members of the prolamin family (Allfam 2011). A number of the prolamin proteins in nuts are allergenic, including 2S albumin in Brazil nuts, cashews, hazelnuts, peanuts, pistachios, soybeans, and walnuts (also see other references listed in Table 2). Cupins are a diverse group of proteins found in bacteria, fungi, animals, and plants; these share one or more double-stranded cupin domains that have been described as “barrel-like” or resembling a “jelly-roll”. Two distinct groups of cupin seed storage proteins are the legumins (11S globulins) and vicilins (7S or 8S globulins) (Allfam 2011) and these make up as much as 70% of the protein of some seeds (Bewley and Black 1994). Allergic cupins have been identified in almonds, Brazil nuts, cashews, hazelnuts, peanuts, pistachios, soybeans, and walnuts (Allergen.org 2014 and other references listed in Table 2). Profilins are a conserved group of proteins found in every eukaryote and in some viruses (Radauer and Breiteneder 2007) and they play key roles in cell movement and signaling (Witke 2004). Plant profilins are abundant in pollen and present in smaller amounts in other plant structures such as fruits and seeds (Radauer et al. 2006). Allergic profilins have been identified in almonds, hazelnuts, and peanuts (Table 2).

Lehrer et al. (2006) reviewed the properties of allergenic proteins to explore whether allergenicity is predictable based on sequence similarity. These authors and others have suggested that evolutionary relatedness of some edible seeds may play a part in allergen cross-sensitivity, although

preliminary experiments and case studies suggest that IgE cross-reactivity can occur between distantly related species. We sought to systematically test whether distantly related edible seeds have similar amino acid sequences in allergenic seed storage proteins. If this were the case, it could explain the IgE cross-reactivity observed between some nuts and would allow physicians to predict which species a patient is most likely to show cross-sensitivity to, given a known allergy.

MATERIALS AND METHODS

Plant Phylogeny

We reviewed evolutionary relationships of plant species that produce nuts and edible seeds from published studies and mapped them onto a summary phylogenetic tree of vascular plants (Soltis et al. 2000; Stevens 2001 onwards; Li et al. 2004).

Literature Review of Allergen Cross-Reactivity

A literature review was conducted to identify published evidence of serum and skin cross-reactivity between allergenic proteins derived from nuts and seeds. We first searched for allergen cross-reactivity studies in the Allergen.org database (2014), which led us to a number of key studies (Ewan 1996; Tariq et al. 1996; Moneret-Vautrin et al. 1998; Willison et al. 2008; Stutius et al. 2010).

Sequence Selection and Alignment

We focused on two of the major allergenic seed protein families, the cupin super family (vicilin and legumin) and prolamins (2S albumin), to test for convergent evolution by comparing gene trees of the protein sequences with estimates of organismal phylogeny. Three separate proteins were chosen for phylogenetic analysis—legumin (11S globulin), vicilin (2S globulin), and 2S albumin. Sequences were downloaded from Allergen.org, Uniprot, and Genbank by searching protein databases for “legumin,” “vicilin,” or “2S albumin,” and by conducting BLAST searches. Our preliminary dataset included all available sequences from database searches. We removed sequences that did not appear to be orthologous in preliminary trees (they did not resolve within the angiosperms), and redundant sequences for clades that were over-represented in the sampling. Protein alignments of the trimmed dataset were generated with the program Multiple Sequence Comparison by Log-Expectation (MUSCLE v3.8.31, Edgar 2004; Jenkins et al. 2005). The alignments were then modified using the program Gblocks as implemented in SEAVIEW (Guoy et al. 2010) to remove hypervariable regions, with the following stringency settings: “Protein,” “Allow small final blocks,” “Allow gaps within blocks,” “Allow less strict flanking regions.” Then, duplicate sequences were removed by examining the alignment using the following criteria: sequences with the same genus/species identifier and the exact same protein sequence were culled so that only one remained; longer sequences were preferentially retained.

Model Selection and Gene Tree Analyses

Model selection for phylogenetic analyses was conducted separately on each protein alignment using the program

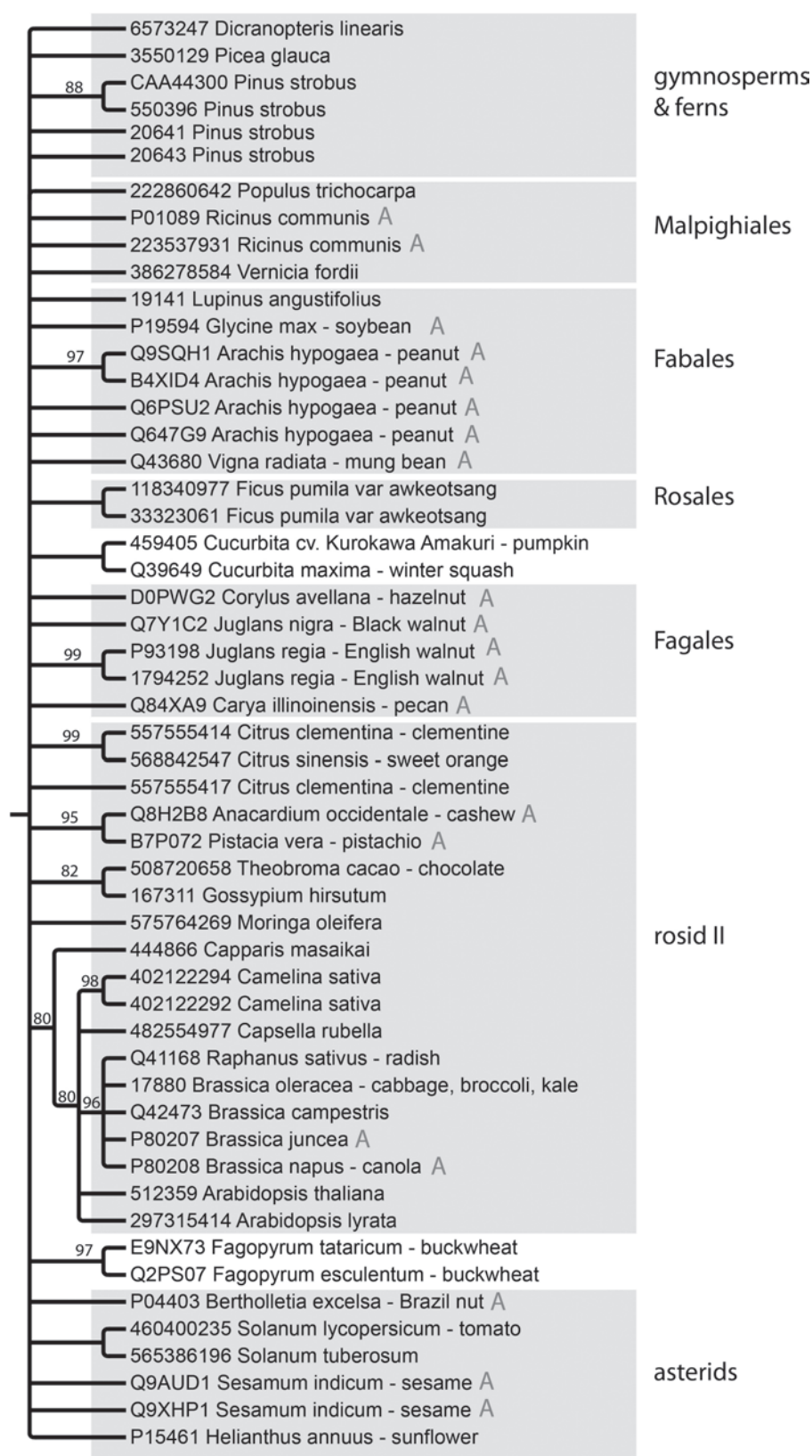


Fig. 2. Estimate of phylogenetic relationships of 2S albumin protein sequences based on a maximum likelihood (ML) analysis in the program RAxML under a JTT + I + gamma model. One fern and five gymnosperm sequences serve as outgroups. ML bootstrap (BS) values (1000 replicates) are printed above branches except when branches had 100% ML BS. Nodes are collapsed if ML BS support was < 80. A = 2S albumin protein allergenic in this species.

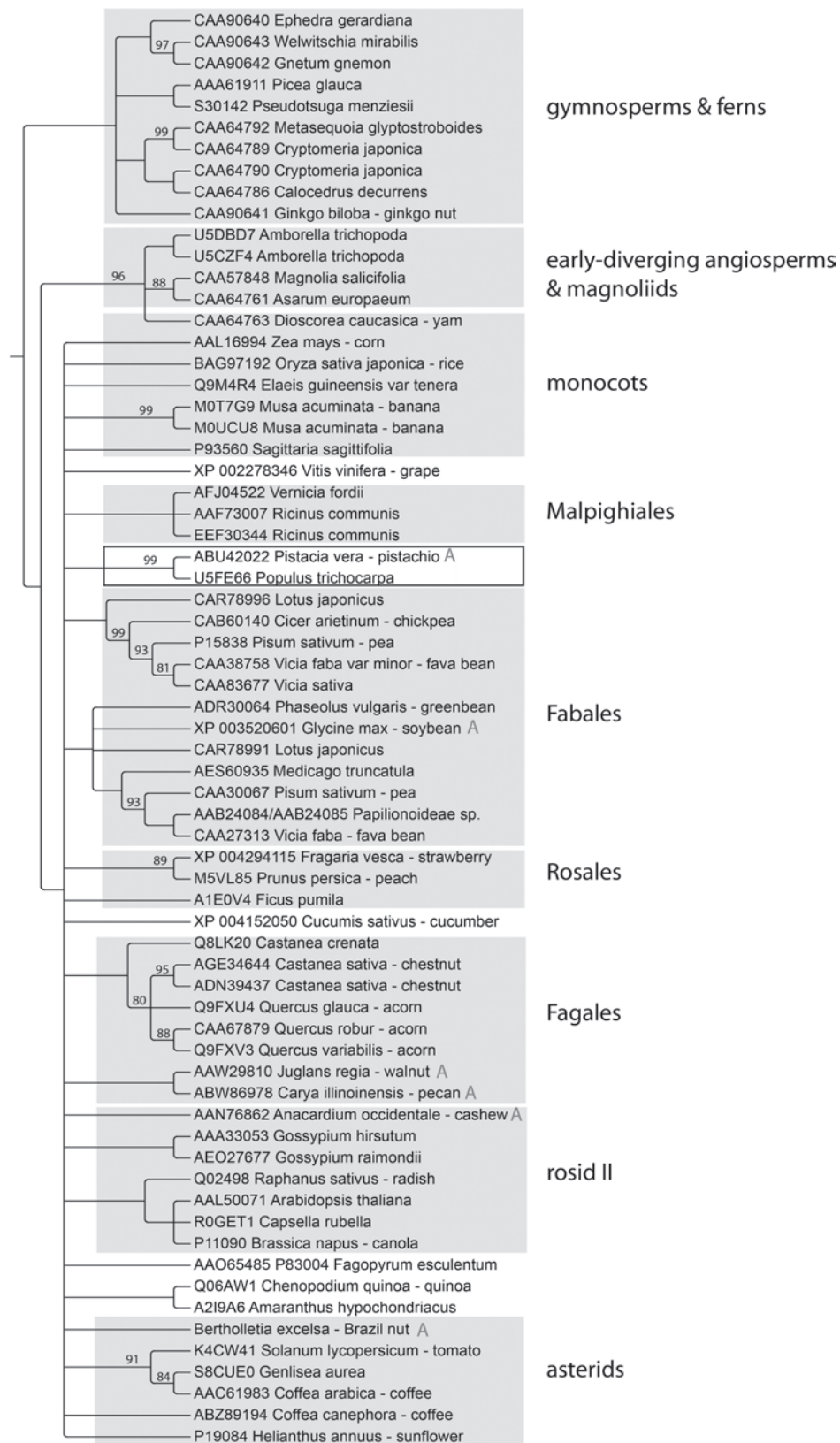


Fig. 3. Estimate of phylogenetic relationships of legumin protein sequences based on a maximum likelihood (ML) analysis in the program RAxML under a LG + I + gamma model. We rooted the tree on the branch leading to (gymnosperms + ferns) + angiosperms. ML bootstrap (BS) values (1000 replicates) are printed above branches except when branches had 100% ML BS. Nodes are collapsed if ML BS support was < 80. A = legumin protein allergenic in this species. The box outlines an unexpectedly close relationship (99% ML BS) between pistachio and poplar (*Populus trichocarpa*).

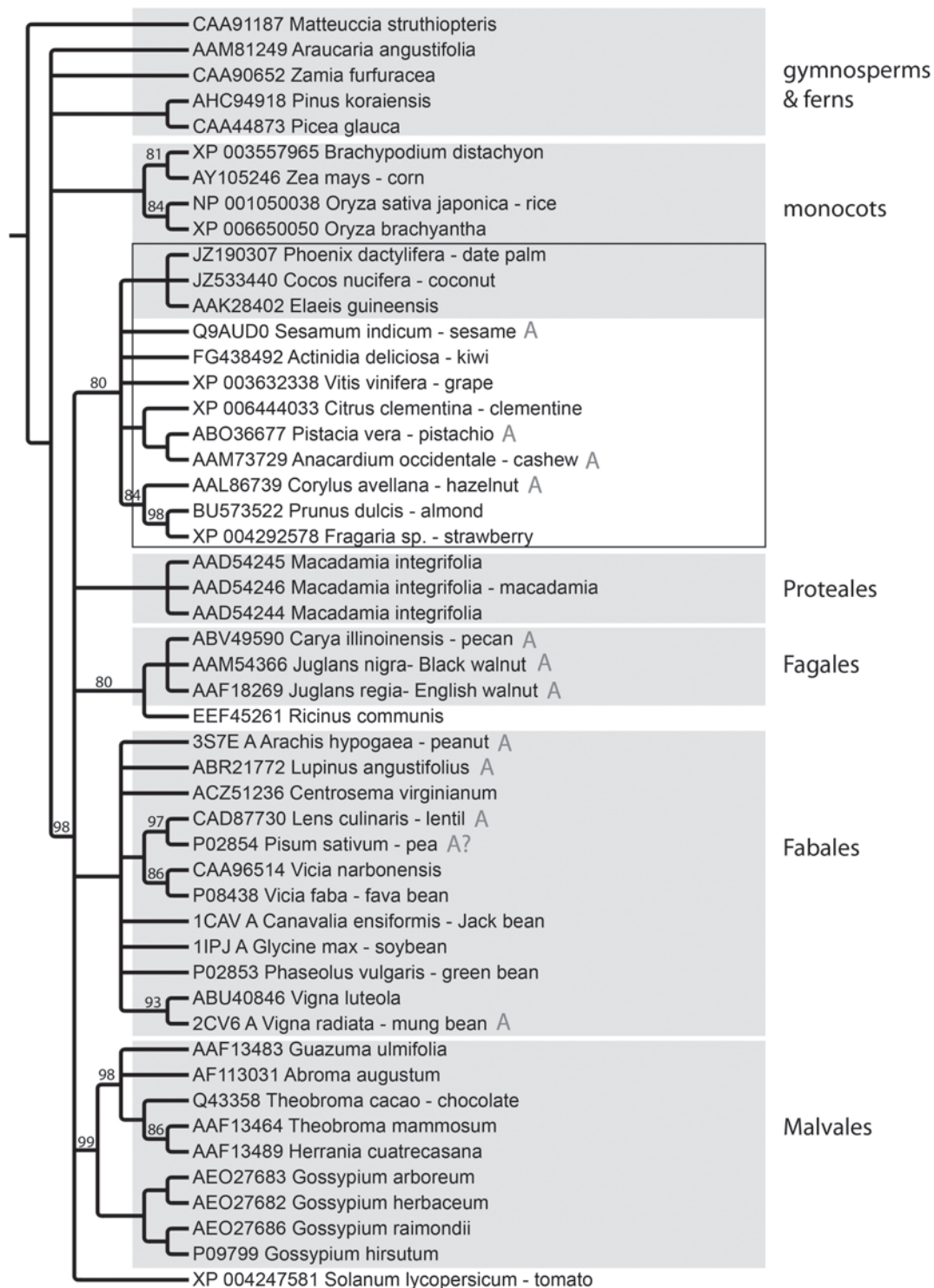


Fig. 4. Estimate of phylogenetic relationships of vicilin protein sequences based on a maximum likelihood (ML) analysis in the program RAxML under an LG + I + gamma model. The vicilin dataset contains one fern (*Matteuccia*), and four gymnosperms (three Pinales and one cycad). We rooted the tree on the branch leading to the fern. ML bootstrap (BS) values (1000 replicates) are printed above branches except when branches had 100% ML BS. Nodes are collapsed if ML BS support was < 80%. A = vicilin protein allergenic in this species. The box outlines an unexpectedly close relationship (80% ML BS) between non-grass monocots (date palm, coconut, and oil palm [*Elaeis*]), kiwi, grape, sesame, citrus, pistachio, cashew (Sapindales), hazelnut (Fagales), and Rosales (almond and strawberry; 98% ML BS).

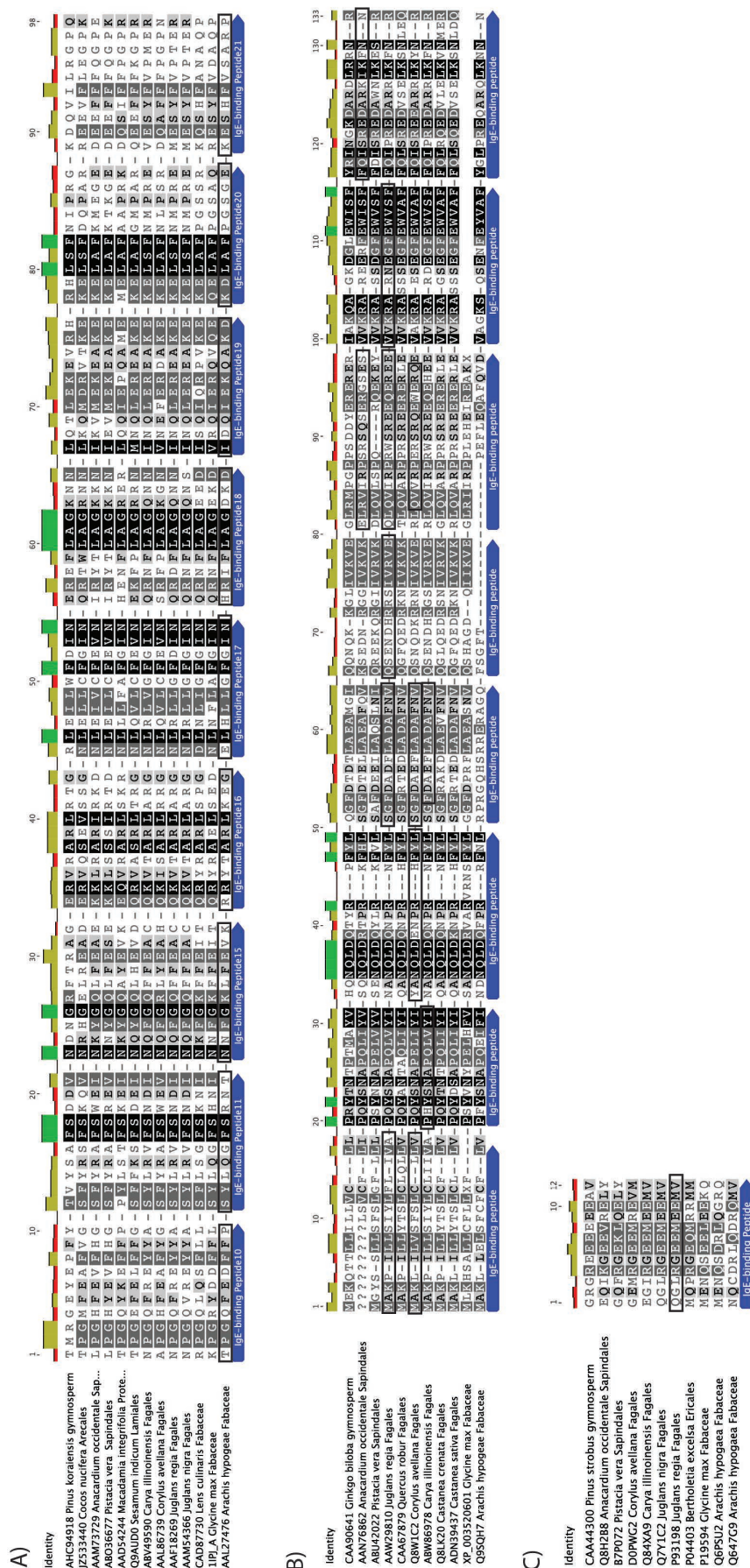


Fig. 5. High levels of amino acid variation in sequences homologous to known Ig-E binding epitopes in A) vicilin peptides, B) legumin, and C) 2S albumin. Black boxes outline sequences that have been experimentally verified as epitopes. The colored graph above the alignment and the shading of each amino acid reflect similarity across the alignment.

Table 4. Sequences included in the study were downloaded from Genbank (GB), the Protein Database (PDB) and Uniprot (UP).

Order	Family	Species name	Legumin	Vicilin	2S albumin
Alismatales	Alismataceae	<i>Sagittaria sagittifolia</i> L.	UP: P93560	–	–
Amborellales	Amborellaceae	<i>Amborella trichopoda</i> Baill.	UP: U5CZF4, U5DBD7	–	–
Araucariales	Araucariaceae	<i>Araucaria angustifolia</i> (Bertol.) Kuntze	–	GB: AAM81249	–
Arecales	Arecaceae	<i>Cocos nucifera</i> L.	–	GB: JZ533440	–
Arecales	Arecaceae	<i>Elaeis guineensis</i> var. <i>tenera</i> Becc.	UP: Q9M4R4	UP: Q9AU64	–
Arecales	Arecaceae	<i>Phoenix dactylifera</i> L.	–	GB: JZ190307	–
Asterales	Asteraceae	<i>Helianthus annuus</i> L.	UP: P19084	–	UP: P15461, GB: AAB41308 GB: XP_002869578
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i> (L.) O’Kane & Al-Shehbaz	–	–	GB: XP_002869578
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i> (L.) Heynh.	GB: AAL50071	–	GB: CAA01132
Brassicales	Brassicaceae	<i>Brassica campestris</i> L.	–	–	UP: Q42473
Brassicales	Brassicaceae	<i>Brassica juncea</i> (L.) Czern.	–	–	UP: P80207
Brassicales	Brassicaceae	<i>Brassica napus</i> L.	UP: P11090	–	UP: P80208
Brassicales	Brassicaceae	<i>Brassica oleracea</i> L.	–	–	GB: CAA46783
Brassicales	Brassicaceae	<i>Camelina sativa</i> (L.) Crantz	–	–	GB: AFQ32283, AFQ32283
Brassicales	Brassicaceae	<i>Capsella rubella</i> Reut.	UP: R0GET1	–	GB: XP_006286272
Brassicales	Brassicaceae	<i>Raphanus sativus</i> L.	UP: Q02498	–	UP: Q41168
Brassicales	Capparaceae	<i>Capparis masakai</i> H. Lév.	–	–	GB: AAB25171
Brassicales	Moringaceae	<i>Moringa oleifera</i> Lam.	–	–	GB: AHG99681
Caryophyllales	Amaranthaceae	<i>Amaranthus</i> <i>hypochondriacus</i> L.	UP: A2I9A6	–	–
Caryophyllales	Amaranthaceae	<i>Chenopodium quinoa</i> Willd.	UP: Q06AW1	–	–
Caryophyllales	Polygonaceae	<i>Fagopyrum esculentum</i> Moench	UP: Q84TC7, P83004	–	UP: Q2PS07
Caryophyllales	Polygonaceae	<i>Fagopyrum tataricum</i> (L.) Gaertn.	–	–	UP: E9NX73
Cucurbitales	Cucurbitaceae	<i>Cucumis sativus</i> L.	GB: XP_004152050	–	–
Cucurbitales	Cucurbitaceae	<i>Cucurbita</i> cv. Kurokawa Amakuri	–	–	GB: BAA03993
Cucurbitales	Cucurbitaceae	<i>Cucurbita maxima</i> Duchesne	–	–	UP: Q39649
Cupressales	Cupressaceae	<i>Calocedrus decurrens</i> (Torr.) Florin	GB: CAA64786	–	–
Cupressales	Cupressaceae	<i>Cryptomeria japonica</i> (Thunb. ex L. f.) D. Don	GB: CAA64789, CAA64790	–	–
Cupressales	Cupressaceae	<i>Metasequoia glyptostroboides</i> Hu & W.C. Cheng	GB: CAA64792	–	–
Cycadales	Zamiaceae	<i>Zamia furfuracea</i> Aiton	–	GB: CAA90652	–
Ephedrales	Ephedraceae	<i>Ephedra gerardiana</i> Wall. ex C.A. Mey.	GB: CAA90640	–	–
Ericales	Actinidiaceae	<i>Actinidia deliciosa</i> (A. Chev.) C.F. Liang & A.R. Ferguson	–	GB: FG438492	–
Ericales	Lecythidaceae	<i>Bertholletia excelsa</i> Bonpl.	AY221641	–	UP: P04403
Fabales	Fabaceae	<i>Arachis hypogaea</i> L.	–	PDB: 3S7E_A	GB: Q6PSU2, Q647G9, UP: B4XID4
Fabales	Fabaceae	<i>Canavalia ensiformis</i> (L.) DC.	–	PDB: 1CAV_A	–
Fabales	Fabaceae	<i>Centrosema virginianum</i> (L.) Benth.	–	GB: ACZ51236	–
Fabales	Fabaceae	<i>Cicer arietinum</i> L.	GB: CAB60140	–	–
Fabales	Fabaceae	<i>Glycine max</i> (L.) Merr.	GB: XP_003520601	PDB: 1IPJ_A	UP: P19594
Fabales	Fabaceae	<i>Lens culinaris</i> Medik.	–	GB: CAD87730	–
Fabales	Fabaceae	<i>Lotus japonicus</i> (Regel) K. Larsen	GB: CAR78996, CAR78991	–	–
Fabales	Fabaceae	<i>Lupinus angustifolius</i> Blanco	–	GB: ABR21772	GB: CAA37598
Fabales	Fabaceae	<i>Medicago truncatula</i> Gaertn.	GB: AES60935	–	–
Fabales	Fabaceae	Papilionoideae indet.	GB: AAB24084, AAB24085	–	–

Table 4. Continued

Order	Family	Species name	Legumin	Vicilin	2S albumin
Fabales	Fabaceae	<i>Phaseolus vulgaris</i> L.	GB: ADR30064	UP: P02853	–
Fabales	Fabaceae	<i>Pisum sativum</i> L.	GB: P15838, CAA30067	GB: P02854	–
Fabales	Fabaceae	<i>Vicia faba</i> L.	GB: CAA38758, CAA27313	UP: P08438	–
Fabales	Fabaceae	<i>Vicia narbonensis</i> L.	–	GB: CAA96514	–
Fabales	Fabaceae	<i>Vicia sativa</i> L.	GB: CAA83677	–	–
Fabales	Fabaceae	<i>Vigna lutea</i> L.	–	GB: ABU40846	–
Fabales	Fabaceae	<i>Vigna radiata</i> L.	–	PDB: 2CV6_A	UP: Q43680
Fagales	Betulaceae	<i>Corylus avellana</i> L.	–	GB: AAL86739	UP: D0PWG2
Fagales	Fagaceae	<i>Castanea crenata</i> Siebold & Zucc.	UP: Q8LK20	–	–
Fagales	Fagaceae	<i>Castanea sativa</i> Mill.	GB: ADN39437, AGE34644	–	–
Fagales	Fagaceae	<i>Quercus glauca</i> Thunb.	UP: Q9FXU4	–	–
Fagales	Fagaceae	<i>Quercus robur</i> L.	GB: CAA67879	–	–
Fagales	Fagaceae	<i>Quercus variabilis</i> Blume	UP: Q9FXV3	–	–
Fagales	Juglandaceae	<i>Carya illinoensis</i> (Wangenh.) K. Koch	GB: ABW86978	GB: ABV49590	UP: Q84XA9
Fagales	Juglandaceae	<i>Juglans nigra</i> L.	–	GB: AAM54366	UP: Q7Y1C2
Fagales	Juglandaceae	<i>Juglans regia</i> L.	GB: AAW29810	GB: AAF18269	UP: P93198
Gentianales	Rubiaceae	<i>Coffea arabica</i> L.	GB: AAC61983	–	–
Gentianales	Rubiaceae	<i>Coffea canephora</i> Pierre ex A. Froehner	GB: ABZ89194	–	–
Ginkgoales	Ginkgoaceae	<i>Ginkgo biloba</i> L.	GB: CAA90641	–	–
Gleicheniales	Gleicheniaceae	<i>Dicranopteris linearis</i> (Burm. f.) Underw.	–	–	GB: AAF17603
Gnetales	Gnetaceae	<i>Gnetum gnemon</i> L.	GB: CAA90642	–	–
Lamiales	Lentibulariaceae	<i>Genlisea aurea</i> A. St.-Hil.	UP: S8CUE0	–	–
Lamiales	Pedaliaceae	<i>Sesamum indicum</i> L.	–	UP: Q9AUD0	UP: Q9AUD1, Q9XHP1
Liliales	Dioscoreaceae	<i>Dioscorea caucasica</i> Lipsky	GB: CAA64763	–	–
Magnoliales	Magnoliaceae	<i>Magnolia salicifolia</i> Maxim.	GB: CAA57848	–	–
Malpighiales	Euphorbiaceae	<i>Ricinus communis</i> L.	GB: AAF73007, EEF30344	GB: EEF45261	UP: P01089, GB: XP_002522847
Malpighiales	Euphorbiaceae	<i>Vernicia fordii</i> (Hemsl.) Airy Shaw	GB: AFJ04522	–	GB: AFJ04524
Malpighiales	Salicaceae	<i>Populus trichocarpa</i> Torr. & A. Gray	UP: U5FE66	–	GB: XP_002317577
Malvales	Malvaceae	<i>Abroma augustum</i> (L.) L. f.	–	GB: AF113031	–
Malvales	Malvaceae	<i>Gossypium arboreum</i> L.	–	GB: AEO27683	–
Malvales	Malvaceae	<i>Gossypium herbaceum</i> L.	–	GB: AEO27682	–
Malvales	Malvaceae	<i>Gossypium hirsutum</i> L.	GB: AAA33053	GB: P09799	GB: AAA33049
Malvales	Malvaceae	<i>Gossypium raimondii</i> Ulbr.	GB: AEO27677	GB: AEO27686	–
Malvales	Malvaceae	<i>Guazuma ulmifolia</i> Lam.	–	GB: AAF13483	–
Malvales	Malvaceae	<i>Herrania cuatrecasana</i> García-Barr.	–	GB: AAF13489	–
Malvales	Malvaceae	<i>Theobroma cacao</i> L.	–	GB: Q43358	GB: EOY12555
Malvales	Malvaceae	<i>Theobroma mammosum</i> Cuatrec. & J. León	–	GB: AAF13464	–
Pinales	Pinaceae	<i>Picea glauca</i> (Moench) Voss	GB: AAA61911	–	GB: AAC34613
Pinales	Pinaceae	<i>Pinus koraiensis</i> Siebold & Zucc.	–	GB: AHC94918	–
Pinales	Pinaceae	<i>Pinus strobus</i> L.	–	–	GB: CAA44298, CAA44301, CAA44299
Pinales	Pinaceae	<i>Pseudotsuga menziesii</i> (Mirb.) Franco	GB: S30142	–	–
Piperales	Aristolochiaceae	<i>Asarum europaeum</i> L.	GB: CAA64761	–	–
Poales	Poaceae	<i>Brachypodium distachyon</i> (L.) P. Beauv.	–	GB: XP_003557965	–
Poales	Poaceae	<i>Oryza brachyantha</i> A. Chev. & Roehr.	–	GB: XP_006650050	–
Poales	Poaceae	<i>Oryza sativa</i> L. subsp. <i>japonica</i> S. Kato	GB: BAG97192	GB: NP_001050038	–

Table 4. Continued

Order	Family	Species name	Legumin	Vicilin	2S albumin
Poales	Poaceae	<i>Zea mays</i> L.	GB: AAL16994	GB: AY105246	–
Polypodiales	Onocleaceae	<i>Matteuccia struthiopteris</i> (L.) Tod.	–	GB: CAA91187	–
Proteales	Proteaceae	<i>Macadamia integrifolia</i> Maiden & Betche	–	GB: AAD54244, AAD54245, AAD54246	–
Rosales	Moraceae	<i>Ficus pumila</i> L.	UP: A1E0V4	–	GB: ABK80756, AAQ07270
Rosales	Rosaceae	<i>Fragaria vesca</i> L. subsp. <i>vesca</i>	GB: XP_004294115	GB: XP_004292578	–
Rosales	Rosaceae	<i>Prunus dulcis</i> (Mill.) D.A. Webb	–	GB: BU573522	–
Rosales	Rosaceae	<i>Prunus persica</i> (L.) Batsch	UP: M5VL85	–	–
Sapindales	Anacardiaceae	<i>Anacardium occidentale</i> L.	GB: AAN76862	GB: AAM73729	UP: Q8H2B8
Sapindales	Anacardiaceae	<i>Pistacia vera</i> L.	GB: ABU42022	GB: ABO36677	UP: B7P072
Sapindales	Rutaceae	<i>Citrus clementina</i> hort. selection	–	GB: XP_006444033	GB: XP_006452191, XP_006452188
Sapindales	Rutaceae	<i>Citrus sinensis</i> (L.) Osbeck	–	–	GB: XP_006475207
Solanales	Solanaceae	<i>Solanum lycopersicum</i> L.	UP: K4CW41	GB: XP_004247581	GB: XP_004245641
Solanales	Solanaceae	<i>Solanum tuberosum</i> L.	–	–	GB: XP_006358919
Vitales	Vitaceae	<i>Vitis vinifera</i> L.	GB: XP_002278346	GB: XP_003632338	–
Welwitschiales	Welwitschiaceae	<i>Welwitschia mirabilis</i> Hook.f.	GB: CAA90643	–	–
Zingiberales	Musaceae	<i>Musa acuminata</i> Colla	UP: M0UCU8, M0T7G9	–	–

ProtTest (Abascal et al. 2005), with the following settings: “Build BioNJ Tree,” “LnL as model selection criteria,” “Optimize tree topology, no.” The best model available for maximum likelihood (ML) analysis in RAxML was selected based on the LnL score (Table 3).

ML phylogenetic analyses were conducted separately on each protein alignment in RAxML (Miller et al. 2010; Stamatakis et al. 2012) using the following settings: “protein,” “estimate proportion of invariable sites GTRGAMMA+I, yes,” “use empirical base frequencies, no.” A protein substitution matrix was chosen for each analysis based on ProtTest results (see above). 1000 bootstrap (BS) replicates were used to estimate node support. Rooting with fern and gymnosperm outgroups was done in the software package FigTree (Bouckaert et al. 2014). Final trees were visualized and edited in TreeGraph2 (Stöver and Miller 2010) and Adobe Illustrator (Adobe Systems Incorporated, San Jose, California).

Gene trees and the best estimate of organismal phylogeny were compared to detect conflicting relationships that were supported by BS values greater than 80%. Incongruence between the gene tree and phylogeny was considered potential evidence for protein sequence convergence. Sequences involved in the conflicting relationships were checked to ensure they were correctly aligned and did not contain ambiguous bases.

Epitope Mapping

The MUSCLE alignment for each protein was annotated with experimentally verified epitopes. The vicilin annotation followed the mapping of IgE binding sites in peanut (*Arachis*) and used the numbering system of Shin et al. (1998). The legumin annotation followed the epitopes of pecan (*Carya*

illinoensis; Sharma et al. 2011) and homologous epitopes documented in the Structural Database of Allergenic Proteins (SDAP; Ivanciuc et al. 2003). The 2S albumin alignment was annotated with the epitope identified in walnut (*Juglans regia*; Robotham et al. 2002).

RESULTS

Plant Phylogeny

Fruit types are not generally conserved within plant orders, although there are some notable exceptions, including the true nuts. Many trees in the order Fagales produce nuts that can be described as indehiscent fruits with a hard covering and a single seed. The most familiar nut in temperate zones may be the acorn that is the fruit of oak trees (*Quercus*), which was a common food in North America only a few hundred years ago. Close relatives to the oak are chestnut (*Castanea*), hazelnut (*Corylus*), pecan (*Carya*), and walnut (*Juglans*). Evolutionary relationships among these species have been difficult to estimate (Manos and Steele 1997; Soltis et al. 2000; Ruiqi et al. 2002; Cook and Crisp 2005; Sauquet et al. 2012), but analyses of DNA sequence data are converging on the topology shown in Fig. 1 (Soltis et al. 2000; Stevens 2001 onwards; Li et al. 2004; Herbert et al. 2006).

Taxa and Sequences Sampled and Final Alignments

The complete listing of sampled taxa and protein GenBank accession numbers can be found in Table 4. The 2S albumin analysis contained 53 taxa and 115 characters. The legumin analysis contained 67 taxa and 337 characters. The vicilin analysis contained 50 taxa and 316 characters.

Table 5. Summary of nut cross-reactivity studies. This table illustrates how little is known about nut and seed cross-reactivity. The rectangles are shaded according to a scale so that a light rectangle indicates no or low evidence for some cross-reactivity and a dark rectangle indicates strong evidence for some cross-reactivity. Many combinations have never been tested, especially for infrequently eaten nuts and several studies have a sample size of $N = 1$. Ewan (1996) tested 62 patients for cross-reactivity to six nuts: peanut, Brazil nut, almond, hazelnut, walnut, and cashew nut, but did not report the species involved in each treatment. Table based on references (Altenbach et al. 1987; Arshad et al. 1991; Fernandez et al. 1995; Tariq et al. 1996; Marinas et al. 1998; Moneret-Vautrin et al. 1998; Sutherland et al. 1999; Teuber and Peterson 1999; Teuber et al. 1999, 2003; Diaz-Perales et al. 2000; Bannon et al. 2001; Poltronieri et al. 2002; Wang et al. 2002; de Leon et al. 2003; Roux et al. 2003; Asero et al. 2004; Lerch et al. 2005; Crespo et al. 2006; Benito et al. 2007; Willison et al. 2008; Ahn et al. 2009; Breiteneder 2009; Garino et al. 2010; Allergen.org 2014).

	Peanut (<i>Arachis hypogaea</i>)	Almond (<i>Prunus dulcis</i>)	Brazil nut (<i>Bertholletia excelsa</i>)	Cashew (<i>Anacardium occidentale</i>)	Hazelnut (<i>Corylus avellana</i>)	English walnut (<i>Juglans regia</i>)
Almond (<i>Prunus dulcis</i>)	51.04% (49/96)	—	—	—	—	—
Brazil nut (<i>Bertholletia excelsa</i>)	28.4% (23/81)	6.66% (1/15)	—	—	—	—
Cashew (<i>Anacardium occidentale</i>)	39.51% (32/81)	—	—	—	—	—
Hazelnut (<i>Corylus avellana</i>)	20.83% (20/96)	28.57% (6/21)	12.5% (2/16)	—	—	—
English walnut (<i>Juglans regia</i>)	17.65% (3/17)	30% (6/20)	100% (1/1)	—	100% (1/1)	—
Macadamia (<i>Macadamia integrifolia</i>)	0% (0/2)	0% (0/2)	0% (0/1)	—	100% (2/2)	0% (0/1)
Pine nut (<i>Pinus edulis</i>)	—	100% (1/1)	—	—	—	—
Pistachio (<i>Pistacia vera</i>)	31% (22/72)	—	—	76.19% (16/21)	—	—
Coconut (<i>Cocos nucifera</i>)	0% (N not reported)	—	—	—	—	100% (2/2)
Pecan (<i>Carya illinoensis</i>)	—	—	—	—	—	100% (1/1)
Sesame (<i>Sesamum indicum</i>)	13% (9/69)	—	—	—	—	—

Gene Trees

The 2S albumin maximum likelihood (ML) estimate was poorly resolved, but exhibits relationships congruent with our current knowledge of plant orders if we consider branches with bootstrap support (BS) > 80 (Fig. 2). Sampled species appear to have 2S albumin sequences most closely related to the 2S albumin sequences of their close relatives or their placement is unresolved in our analysis.

Similar to the 2S albumin tree, relationships with BS > 80 in the legumin tree (Fig. 3) reflect our current understanding of relationships of plant orders, with a minor exception. Poplar (*Populus trichocarpa*) is recovered outside the Malpighiales clade, and in our tree it is sister to pistachio (Sapindales) legumin with 99 ML BS support (Fig. 3, box).

The majority of the relationships in the vicilin tree (Fig. 4) conform to our current understanding of plant species phylogeny, with some notable exceptions. The main Fagales clade contains a strongly supported (100 ML BS) clade of pecan + walnut with an unexpected sister relationship with castor oil plant (*Ricinus communis*), albeit with only moderate support (80 ML BS). Tomato (*Solanum*, Solanales) is unresolved in the vicilin tree, although we expected it would be sister with sesame (*Sesamum*, Lamiales) and closely related to kiwi (*Actinidia*, Ericales) based on current estimates of plant phylogeny. Instead, kiwi and sesame unexpectedly form a group with non-grass monocots (date palm [*Phoenix*], coconut [*Cocos*] and oil palm [*Elaeis*]), grape (*Vitis*), a well-supported (100 ML BS) clade of citrus, pistachio, and cashew (Sapindales), almond and strawberry (Rosales; 98 ML BS), and hazelnut (*Corylus*, Fagales; 80 ML BS; Fig. 4, box). Interestingly, hazelnut is sister to almond and strawberry with moderate support (84 ML BS), instead of forming a relationship with other Fagales in the dataset.

DISCUSSION

Cross-Reactivity is not Associated with Sequence Similarity

Here, we sought to systematically test whether similarity in the amino acid sequences of seed storage proteins is congruent with the cross-reactivity reported in the literature between distantly related nuts. We conducted independent phylogenetic analyses of three allergenic proteins and found little evidence for sequence convergence. In other words, we did not recover evidence that cross-reactive species that are distantly related have 2S albumin, vicilin, or legumin proteins with similar sequences. The trees we recovered generally lacked strong statistical support (BS values), but the majority of supported relationships do not contradict our current best estimate of evolutionary history and therefore do not suggest convergence between distantly-related allergenic proteins at the level of amino acid sequences.

Our results suggest that cross-reactivity reported in the literature may be difficult to predict and may be the result of factors unrelated to amino acid sequences, as other authors have suggested. Wang et al. (2002) aligned a newly identified vicilin gene in cashews (Ana o 1) with the peanut vicilin gene (Ara h 1). They found few conserved amino acids, even within IgE-binding epitopes, and hypothesized that protein sequence similarity is not necessarily predictive of allergenicity. We also found that the amino acid sequence was highly variable in areas of the alignment that have been identified as IgE-binding sites in some species (Fig. 5). Barre et al. (2005) showed that there is little structural homology between Brazil nut, peanut, pecan, or walnut 2S albumin proteins, but the authors did find a similar three-dimensional protein structure in walnut and pecan and these species are close relatives (Barre et al. 2005). Thus, factors such as structural homology may play a role in

cross-reactivity, though additional comparisons are needed to investigate this hypothesis.

Pine Nuts and Ginkgo Nuts

Pine nuts and *Ginkgo* nuts are both gymnosperms, but their exact relationship to one another is not known (Mathews 2009; Ran et al. 2010; Yang et al. 2012). Pine nuts are the seeds of approximately 12 *Pinus* species that are eaten raw, roasted or in cooked dishes (Rosengarten 1984). The seed of the *Ginkgo* tree (*Ginkgo biloba*) is boiled, roasted, and salted before it is eaten. A single near-fatal reaction (Beyer et al. 1998) and four systemic reactions (Nielsen 1990; Roux et al. 1998) have been reported after pine nut ingestion. In one patient, pine nut showed cross-reactivity to almonds (Marinas et al. 1998; Table 5), but cross-reactivity to *Ginkgo* was not tested. As far as we are aware, *Ginkgo* nuts have only been tested for contact dermatitis (Lepoittevin et al. 1989). We sampled *Ginkgo* only for the legumin dataset. This species is currently not an important food source, but is commonly used as an herbal remedy. It would be interesting to study whether the same seed-storage proteins are present in these gymnosperms, which proteins are allergenic, and whether they are cross-reactive.

Palm Nuts

A number of seeds from palm species (monocots in the Arecaceae family) are roasted and eaten as nuts. These include the date palm (*Phoenix dactylifera*) and the ivory nut palm (*Phytelephas aequatorialis*) (Rosengarten 1984). Coconuts (*Cocos nucifera*, Arecaceae) are also considered nuts by some, and the coconut milk or meat that is eaten is the seed nutritive tissue. In two patients, cross-reactivity to tree nuts has been observed (Teuber and Peterson 1999). One patient who showed anaphylaxis to coconut also demonstrated cross-reactivity to hazelnut (Nguyen et al. 2004). A retrospective chart review of 231 patients who underwent skin prick tests to determine sensitization demonstrated that coconut allergy was not more common in patients with sensitization or allergy to peanuts or tree nuts than in those without peanut or tree nut allergies (Stutius et al. 2010).

Peanuts, Soy, Almonds, Hemp Nuts

These nuts are in the rosid I clade of eudicot plants. We recovered a bean and pea (Fabales) clade in the legumin and vicilin trees, and the 2S albumin tree is congruent with a Fabales clade. Peanuts are the most widely tested species for cross-reactivity with other seeds (Table 5): peanuts exhibited 51% cross-reactivity with almonds, a close relative in the Rosales (Tariq et al. 1996; Moneret-Vautrin et al. 1998), 20.8% cross-reactivity with hazelnuts (Tariq et al. 1996; Moneret-Vautrin et al. 1998), and 17.6% cross-reactivity with English walnuts (Tariq et al. 1996; de Leon et al. 2003). They also exhibit similar cross-reactivity with more distant relatives such as Brazil nuts (28.4%; Moneret-Vautrin et al. 1998; de Leon et al. 2007), cashews (39.5%; Moneret-Vautrin et al. 1998; de Leon et al. 2003), pistachios (31%; Moneret-Vautrin et al. 1998), and sesame seeds (13%; Stutius et al. 2010). de Leon et al. (2007) tested sera from three subjects for IgE cross-reactivity between a 2S albumin peanut extract and almond,

Brazil nut, cashew, and hazelnut extracts and demonstrated a strong cross-reaction between peanut 2S albumin and roasted almond and a weaker interaction between peanut 2S albumin and raw Brazil nut. The allergens that have been identified in peanuts are vicilin, 2S albumin, and lipid-transfer proteins (Allergen.org 2014). We sampled the vicilin and 2S albumin genes for peanut and found no evidence of amino acid sequence convergence between species, including closely-related species like mung bean and lentil, which are also allergenic for these proteins. The edible Tahitian chestnut or mape (*Inocarpus fagifer*) is also a close relative to peanuts, but we found no information about allergies to this species.

Almonds (*Prunus dulcis*, Rosaceae), hemp nuts (*Cannabis sativa*, Cannabaceae), and breadnuts (the common name for two distinct species, *Artocarpus camansi* in Polynesia and *Brosimum alicastrum* in the neotropics) are in the Rosales order. We recovered a small Rosales clade in all of the trees, but in the vicilin tree, almond and strawberry (both in the Rosaceae) were unexpectedly sister to hazelnut (Fagales) on a branch with moderate support (84 ML BS; Fig. 4). Vicilin has been identified as an allergen in hazelnut (Roux et al. 2003), but not in almond or strawberry. Hazelnut and almond exhibited cross-reactivity in 6 out of 21 patients (28.6%; Tariq et al. 1996; Poltronieri et al. 2002) and the same level of reactivity was found in those studies between almond and walnut, yet walnut is sister to pecan in our vicilin tree. Almonds are the most commonly consumed tree nut in the United States (Roux et al. 2003) and their identified allergens include 2S albumin (Poltronieri et al. 2002), legumin (Willison et al. 2011), lipid-transfer protein, and profilin (Allergen.org 2014). Hemp nuts (*Cannabis sativa*, Rosales) were not sampled in our study, but lipid transfer proteins have been identified from the nuts and allergies to hemp nuts have been reported (Ebo et al. 2013). To our knowledge, breadnuts have not been tested for allergens.

In the vicilin tree (Fig. 4) the castor oil plant, *Ricinus* (Euphorbiaceae, Malpighiales), is sister to pecan and walnut (Fagales). Vicilin has been identified as an allergen in walnut, but not in pecan or castor oil plant. Salcedo et al. (2001) found cross-reactivity between rubber latex allergens (natural latex is derived from the rubber tree *Hevea* and both *Hevea* and *Ricinus* are in the Euphorbiaceae), but this was thought to have been due to the presence of profilin proteins. We did not include profilins in this study because their amino acid sequences are highly conserved across plants (Vieths et al. 2002; Radauer et al. 2006).

The tropical candlenut (or kukui, *Aleurites moluccana*) and kluwak nut (*Pangium edule*) are sister to the clade containing the Fagales (true nuts), the Rosales, and the Fabales. We found no sequence data or information about allergic reactions to these species.

Fagales

Acorns, beechnuts, butternuts, chestnuts, hazelnuts, hickory nuts, pecans, and walnuts are closely related fruits in the Fagales order. We recovered tree topologies that are congruent with the current best estimate of evolutionary history in the Fagales, except for the placement of hazelnut in the vicilin tree. Walnut and pecan are close relatives in the Juglandaceae family. Hazelnut is in the Betulaceae family, while chestnut,

oak, and beechnut are close relatives in the Fagaceae family (Stevens 2001 onwards; Li et al. 2004). 100% cross-reactivity has been reported between walnut and pecan (Teuber et al. 2003), walnut and hazelnut (Asero et al. 2004), and other nuts in this order (Bock and Atkins 1989; Ewan 1996; Sicherer et al. 1998; Teuber et al. 2003). These reports are usually restricted to a single patient with multiple allergies and it is unclear if the sensitivity is caused by cross-reactivity to a protein found in both nuts or to co-sensitivity. Additionally, Barre et al. (2005) found that the epitope sequences and overall protein structures are highly similar in walnut and pecan 2S albumin. Given how closely related these species are it seems reasonable to avoid eating any nuts or seeds from plants in the Fagales order if there are signs of allergy to one species.

Pistachios and Cashews

These are closely related species in the Anacardiaceae family (Sapindales, rosid II) and are also closely related to pili nut (*Canarium ovatum*, Burseraceae), lychee (*Litchi chinensis*, Sapindaceae), whose seed is sometimes eaten as a nut, and kola nut (*Cola acuminata*, Malvales). Poison ivy (*Toxicodendron*, Anacardiaceae) and mangos (Anacardiaceae) are also close relatives that commonly cause allergic reactions. A more distant relative, but also in the rosid II clade, is the water chestnut (*Trapa natans*, Myrtales). Water chestnuts were once a major source of starch for central Europeans (Rosengarten 1984). Species in the genus *Trapa* are now primarily eaten in China, Japan, and Korea (Rosengarten 1984), but we found no information about them as allergens.

Pistachio and cashew exhibit 76% cross-reactivity in serum analyses (Fernandez et al. 1995; de Leon et al. 2003), but they have not been tested for cross-reactivity with nuts other than peanuts. Vicilin, legumin, and 2S albumin have been identified as allergens in pistachio and cashew (Wang et al. 2002; Allergen.org 2014) and we found a close relationship between their 2S albumin and vicilin protein sequences, but in the legumin tree pistachio is sister to poplar (Malpighiales, rosid I) and cashew is unresolved. It is not clear to us why the pistachio legumin sequence would be similar to poplar legumin, a common pollen allergen, but this relationship deserves further study.

Distant Relatives to All Other Nuts: Macadamia Nuts, Brazil Nuts, Sesame and Sunflower Seeds

Macadamia nuts (*M. integrifolia* or *M. tetraphylla*) are in an early-diverging lineage of the eudicots (Proteales) and are only distantly related to other commonly eaten seeds. We sampled three forms of vicilin from macadamia and they formed a clade. There have been reported cases of allergic reaction after ingestion of macadamia nuts, but the allergenic protein has yet to be identified (Roux et al. 2003 and references therein). Studies involving a total of two patients found that both patients were cross-reactive to hazelnut and macadamia, but cross-reactivity did not extend to almond, Brazil nut, peanut or walnut (de Leon et al. 2003; Lerch et al. 2005). We found no relationship between hazelnut and macadamia vicilin sequences, but one of the other allergenic proteins in hazelnut may be responsible for the cross-reactivity and this relationship deserves more study.

Brazil nut (*Bertholletia excelsa*) is classified in the Ericales order of the asterids, along with paradise nuts (*Lecythis zabucajo*) and shea nuts (*Vitellaria paradoxa*), two species rarely eaten in the US or Europe. We sampled the allergenic Brazil nut 2S albumin and legumin sequences and Brazil nut was unresolved in both trees. Brazil nut proteins elicited a cross-reaction with peanuts in 28.4% of patients (Moneret-Vautrin et al. 1998; de Leon et al. 2007), with hazelnuts in 12.5% of patients (Tariq et al. 1996), with walnuts in a single patient (Arshad et al. 1991), with almonds in 6.7% of patients (Tariq et al. 1996), and didn't cross-react with macadamia in the single patient tested (Lerch et al. 2005) (Table 5). In most of these tests for cross-reactivity the patient sample size was quite low and we can offer no explanation for the observed pattern of cross-reactivity given the evolutionary distance between Ericales and these other nuts.

Sesame seeds (*Sesamum indicum*, Pedaliaceae) are one of the few edible seeds besides chia (*Salvia* spp.) in the mint order (Lamiales) of plants. Sesame seeds are mostly pressed for oil, but are also used in breads, sauces (mole), and pastes (tahini). Sesame allergens include vicilin, legumin, 2S albumin, and oleosin (Pastorello et al. 2001; Breiteneder 2009). We sampled several accessions of sesame 2S albumin, but its relationship to other species was unresolved (Fig. 2). Sesame is a major cause of food allergy in some countries (Dalal et al. 2002; Osborne et al. 2011) and may be increasingly common in the US (Zuidmeer et al. 2008). Stutius et al. (2010) found that only 9 out of 69 patients with peanut allergies were also sensitive to sesame (Table 5).

Sunflowers are in the Asterales order of plants and are a commonly eaten seed in the United States. Several reports have been published on sensitization and allergy to sunflower seeds (Noyes et al. 1979; Axelsson et al. 1994) and sensitivity seems to be provoked by birch or mugwort pollen (Vieths et al. 2002). Sunflower seed lipid transfer protein has been identified as an allergen (Yagami 2010), while profilin in this species has been shown to be non-allergenic (Asturias et al. 1998). In one case, a patient with sunflower seed allergy was not reactive to other plants in the Asterales (mugwort, ragweed, dandelion) or to nuts (almonds, Brazil nuts, peanuts; Yagami 2010). There are no commonly eaten edible seeds that are closely related to sunflowers.

Study Limitations

Our study sampled three proteins from the five known, common allergenic proteins (Table 2), and it is likely that researchers have not yet identified all of the allergenic proteins in nuts and seeds. Many proteins are present in small amounts and may be difficult to isolate and identify, but next-generation sequencing may quickly increase the pace of discovery of allergenic proteins. In particular, transcriptomes of nuts and seeds may yield protein sequences of potential allergens that will allow researchers a much larger database from which they can compare allergenic and non-allergenic proteins (Radauer et al. 2008). This will allow for more robust tests of proteins and protein characteristics (such as tertiary structure) that are shared by allergens.

We have referred to cross-reactivity assays such as skin-prick tests and IgE binding tests as evidence for the immune system's reaction to a protein after sensitization by exposure to

the protein from another source. However, these assays are problematic, in that their results “do not always correlate with clinical reactivity” (Burks et al. 2012). There are many possible explanations for why cross-reactivity assays are not robust, including that a person may be genetically predisposed to multiple allergies (Sicherer 2002; Morafo et al. 2003; Kulis et al. 2011) or sensitive to many foods (Latham et al. 2003), without proteins from those foods exhibiting cross-reactivity. Food allergies seem to be complicated responses that are partly genetic, partly caused by exposure, and partly caused by the allergenic potential of certain proteins (reviewed in Berin and Sampson 2013). To our knowledge, there have not been nut and seed cross-reactivity studies such as skin prick tests conducted on large patient populations (100s of participants) to allow for statistical tests of correlation. Moneret-Vautin et al.’s (1998) work on peanut cross-reactivity with tree nuts is the exception, with 74 participants tested for cross-reactivity between peanut and five tree nut species, but cross-reactivity between the tree nut species was not tested. Additional cross-reactivity studies are necessary in order to better understand this phenomenon and its underlying causes, if these causes are indeed unrelated to amino acid similarities.

EVOLUTIONARY MEDICINE: USING KNOWLEDGE OF PLANT RELATIONSHIPS TO INFORM ALLERGY STUDIES

In this study, we used phylogenetic analyses to test for similarity in allergenic proteins from edible nuts and seeds across seed-bearing plants. We demonstrated what a number of smaller, non-evolutionary studies have already suggested: that sequence data alone does not explain the multiple nut allergies experienced by some patients. Our data are consistent with the idea that there are structural or chemical features of allergenic proteins that sensitize the immune system and can cause cross-reactions; those features, however, are currently unidentified. Importantly, our method of testing for sequence similarity by comparing new evidence to plant species relationships can be extended to test structural, chemical, and other characters of the proteins.

With this in mind, future research on nut and seed allergens may want to consider an evolutionary framework to generate null hypotheses about cross-reactivity between edible seeds, as such an approach allows for hypothesis testing. For example, we can hypothesize that cross-reactivity may be more common between closely related species in the Fagales order, based on their close evolutionary relationship, and this hypothesis can be tested with in vitro tests between proteins extracted from Fagales nuts listed in Table 1 and other common nuts. If walnuts show higher cross-reactivity rates with distantly related Brazil nuts than their close relatives (as reported from a single patient, Table 3), then the structural and chemical characteristics of proteins in walnuts and Brazil nuts can be studied to identify shared characteristics. Additionally, plants of no culinary value that are closely related to species with common allergens can be studied to understand whether there are characteristics that predispose these proteins to be allergenic. Such an approach would require additional data on the prevalence of cross-reactivity to adequately test which nut species and proteins are cross-reactive at a statistically significant rate. Ultimately, a better understanding of the evolutionary history of plant seed proteins may shed light on

the numerous, unanswered scientific questions regarding tree nut allergen cross-reactivity. We advocate that an evolutionary approach to such unanswered questions could be both novel and fruitful.

ACKNOWLEDGMENTS

We thank the CIPRES web portal for access to computing power and software. Several anonymous reviewers made helpful suggestions to earlier drafts of the manuscript. Fisher’s postdoctoral salary was funded by Rancho Santa Ana Botanic Garden in Claremont, California. The authors also thank Deborah A. Fraser, PhD, for her review of the glossary.

LITERATURE CITED

- ABASCAL, F., R. ZARDOYA, AND D. POSADA. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**: 2104–2105.
- AHN, K., L. BARDINA, G. GRISHINA, K. BEYER, AND H. A. SAMPSON. 2009. Identification of two pistachio allergens, Pis v 1 and Pis v 2, belonging to the 2S albumin and 11S globulin family. *Clinical & Experimental Allergy* **39**: 926–934.
- ALLERGEN.ORG. 2014. Allergen Nomenclature—WHO/IUIS Allergen Nomenclature Sub-Committee. <http://www.allergen.org> (Mar 2015).
- ALLFAM: DATABASE OF ALLERGEN FAMILIES. 2011. <http://www.meduni-wien.ac.at/allergens/allfam/> (March 2015).
- ALTENBACH, S. B., K. W. PEARSON, F. W. LEUNG, AND S. S. SUN. 1987. Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptionally rich in methionine. *Pl. Molec. Biol.* **8**: 239–250.
- ANDERSON, J. A., R. J. EFFERTZ, J. D. FARIS, L. J. FRANCL, S. W. MEINHARDT, AND B. S. GILL. 1999. Genetic analysis of sensitivity to a *Pyrenophora tritici-repentis* necrosis-inducing toxin in durum and common wheat. *Phytopathology* **89**: 293–297.
- ARSHAD, S. H., E. MALMBERG, K. KRAFF, AND D. W. HIDE. 1991. Clinical and immunological characteristics of Brazil nut allergy. *Clinical & Experimental Allergy* **21**: 373–376.
- ASERO, R., G. MISTRELLO, D. RONCAROLO, AND S. AMATO. 2004. Walnut-induced anaphylaxis with cross-reactivity to hazelnut and Brazil nut. *Journal of Allergy and Clinical Immunology* **113**: 358–360.
- ASTURIAS, J. A., M. C. ARILLA, N. GOMEZ-BAYON, M. AGUIRRE, A. MARTINEZ, R. PALACIOS, AND J. MARTINEZ. 1998. Cloning and immunological characterization of the allergen Hel a 2 (profilin) from sunflower pollen. *Molecular Immunology* **35**: 469–478.
- AXELSSON, I. G., E. IHRE, AND O. ZETTERSTROM. 1994. Anaphylactic reactions to sunflower seed. *Allergy* **49**: 517–520.
- BANNON, G. A., G. COCKRELL, C. CONNAUGHTON, C. M. WEST, R. HELM, J. S. STANLEY, N. KING, P. RABJOHN, H. A. SAMPSON, AND A. W. BURKS. 2001. Engineering, characterization and in vitro efficacy of the major peanut allergens for use in immunotherapy. *International Archives of Allergy and Immunology* **124**: 70–72.
- BARRE, A., J.-P. BORGES, R. CULERRIER, AND P. ROUGE. 2005. Homology modelling of the major peanut allergen Ara h 2 and surface mapping of IgE-binding epitopes. *Immunology Letters* **100**: 153–158.
- BENITO, C., E. GONZALEZ-MANCEBO, M. D. DE DURANA, R. M. TOLON, AND M. FERNANDEZ-RIVAS. 2007. Identification of a 7S globulin as a novel coconut allergen. *Annals of Allergy, Asthma & Immunology* **98**: 580–584.
- BERIN, M. C. AND H. A. SAMPSON. 2013. Food allergy: an enigmatic epidemic. *Trends in Immunology* **34**: 390–397.
- BEWLEY, J. D. AND M. BLACK. 1994. Seeds: physiology of development and germination. Plenum Press, New York.
- BEYER, A. V., H. GALL, AND R. U. PETER. 1998. Anaphylaxis to pine nuts. *Allergy* **53**: 1227–1228.

- BEYER, K., G. GRISHINA, L. BARDINA, A. GRISHIN, AND H. A. SAMPSON. 2002. Identification of an 11S globulin as a major hazelnut food allergen in hazelnut-induced systemic reactions. *Journal of Allergy and Clinical Immunology* **110**: 517–523.
- BOCK, S. A. AND F. M. ATKINS. 1989. The natural history of peanut allergy. *Journal of Allergy and Clinical Immunology* **83**: 900–904.
- , A. MUNOZ-FURLONG, AND H. A. SAMPSON. 2001. Fatalities due to anaphylactic reactions to foods. *Journal of Allergy and Clinical Immunology* **107**: 191–193.
- BOUCKAERT, R., J. HELED, D. KÜHNERT, T. VAUGHAN, C.-H. WU, D. XIE, M. A. SUCHARD, A. RAMBAUT, AND A. J. DRUMMOND. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLOS Computational Biology* **10**: e1003537.
- BREITENEDER, H. 2009. The classification of plant food allergens. *Allergologie* **32**: 375–382.
- AND C. RADAUER. 2004. A classification of plant food allergens. *Journal of Allergy and Clinical Immunology* **113**: 821–830.
- BROUK, B. 1975. Plants consumed by man. Academic Press, New York.
- BURKS, A. W., M. TANG, S. SICHERER, A. MURARO, P. A. EIGENMANN, M. EBISAWA, A. FIOCCHI, W. CHIANG, K. BEYER, R. WOOD, J. HOURIHANE, S. M. JONES, G. LACK, AND H. A. SAMPSON. 2012. ICON: food allergy. *Journal of Allergy and Clinical Immunology* **129**: 906–920.
- BURKS, W., G. COCKRELL, S. STANLEY, R. M. HELM, AND G. A. BANNON. 1995. Isolation, identification, and characterization of clones encoding antigens responsible for peanut hypersensitivity. *International Archives of Allergy and Immunology* **107**: 248–250.
- CAREY, A., S. M. POULOSE, AND B. SHUKITT-HALE. 2012. The beneficial effects of tree nuts on the aging brain. *Nutrition and Aging* **1**: 55–67.
- CIANFERONI, A. AND A. MURARO. 2012. Food-induced anaphylaxis. *Immunology & Allergy Clinics* **32**: 165–195.
- CLARKE, B. C., M. HOBBS, D. SKYLAS, AND R. APPELS. 2000. Genes active in developing wheat endosperm. *Functional & Integrative Genomics* **1**: 44–55.
- CODINA, R. M., E. CALDERÓN, R. F. LOCKEY, E. FERNÁNDEZ-CALDAS, AND R. RAMA. 1997a. Specific immunoglobulins to soybean hull allergens in soybean asthma. *CHEST* **111**: 75.
- CODINA, R., R. F. LOCKEY, E. FERNÁNDEZ-CALDAS, AND R. RAMA. 1997b. Purification and characterization of a soybean hull allergen responsible for the Barcelona asthma outbreaks. II. Purification and sequencing of the Gly m 2 allergen. *Clinical & Experimental Allergy* **27**: 424–430.
- COOK, L. G. AND M. D. CRISP. 2005. Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. *Proc. Roy. Soc. Biol. Sci. Ser. B* **272**: 2535–2544.
- CRESPO, J. F., J. M. JAMES, C. FERNÁNDEZ-RODRIGUEZ, AND J. RODRIGUEZ. 2006. Food allergy: nuts and tree nuts. *British Journal of Nutrition* **96**, Suppl. 2: S95–102.
- CROWELL, D. N., M. E. JOHN, D. RUSSELL, AND R. M. AMASINO. 1992. Characterization of a stress-induced, developmentally regulated gene family from soybean. *Pl. Molec. Biol.* **18**: 459–466.
- DALAL, I., I. BINSON, R. REIFEN, Z. AMITAI, T. SHOHAT, S. RAHMANI, A. LEVINE, A. BALLIN, AND E. SOMEKH. 2002. Food allergy is a matter of geography after all: sesame as a major cause of severe IgE-mediated food allergic reactions among infants and young children in Israel. *Allergy* **57**: 362–365.
- DE LEON, M. P., A. C. DREW, I. N. GLASPOLE, C. SUPHIOGLU, R. E. O'HEHIR, AND J. M. ROLLAND. 2007. IgE cross-reactivity between the major peanut allergen Ara h 2 and tree nut allergens. *Molecular Immunology* **44**: 463–471.
- , I. N. GLASPOLE, A. C. DREW, J. M. ROLLAND, R. E. O'HEHIR, AND C. SUPHIOGLU. 2003. Immunological analysis of allergenic cross-reactivity between peanut and tree nuts. *Clinical Allergy* **33**: 1273–1280.
- DIAZ-PERALES, A., M. LOMBARDEO, R. SANCHEZ-MONGE, F. J. GARCIA-SELLES, M. PERNAS, M. FERNÁNDEZ-RIVAS, D. BARBER, AND G. SALCEDO. 2000. Lipid-transfer proteins as potential plant panallergens: cross-reactivity among proteins of *Artemisia* pollen, *Castanea* nut and Rosaceae fruits, with different IgE-binding capacities. *Clinical & Experimental Allergy* **30**: 1403–1410.
- EBISAWA, M., P. BROSTEDT, S. SJÖLANDER, S. SATO, M. P. BORRES, AND K. ITO. 2013. Gly m 2S albumin is a major allergen with a high diagnostic value in soybean-allergic children. *Journal of Allergy and Clinical Immunology* **132**: 976–978.
- EBO, D. G., S. SWERTS, V. SABATO, M. M. HAGENDORENS, C. H. BRIDTS, P. G. JORENS, AND L. S. DE CLERCK. 2013. New food allergies in a European non-Mediterranean region: is *Cannabis sativa* to blame? *International Archives of Allergy and Immunology* **161**: 220–228.
- EDGAR, R. C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113.
- EWAN, P. W. 1996. Clinical study of peanut and nut allergy in 62 consecutive patients: new features and associations. *BMJ* **312**: 1074–1078.
- FERNÁNDEZ, C., A. FIANDOR, A. MARTINEZ-GARATE, AND J. MARTINEZ QUESADA. 1995. Allergy to pistachio: crossreactivity between pistachio nut and other Anacardiaceae. *Clinical & Experimental Allergy* **25**: 1254–1259.
- FRASER, G. E., J. SABATE, W. L. BEESON, AND T. M. STRAHAN. 1992. A possible protective effect of nut consumption on risk of coronary heart disease. The Adventist Health Study. *Archives of Internal Medicine* **152**: 1416–1424.
- GARG, M. L., R. J. BLAKE, AND R. B. H. WILLS. 2003. Macadamia nut consumption lowers plasma total and LDL cholesterol levels in hypercholesterolemic men. *Journal of Nutrition* **133**: 1060–1063.
- GARINO, C. L., J. ZUIDMEER, J. MARSH, M. LOVEGROVE, S. MORATI, P. VERSTEEG, P. SCHILTE, P. SHEWRY, M. ARLORIO, AND R. VAN REE. 2010. Isolation, cloning, and characterization of the 2S albumin: a new allergen from hazelnut. *Molecular Nutrition & Food Research* **54**: 1257–1265.
- GUOY, M., S. GUINDON, AND O. GASCUEL. 2010. Seaview version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molec. Biol. Evol.* **27**: 221–224.
- HARRIS, J. G. AND M. W. HARRIS. 2001. Plant identification terminology. Spring Lake Publishing, Payson, Utah.
- HEDGES, S. B. 2009. Vertebrates (Vertebrata), pp. 309–314. In: S. B. Hedges and S. Kumar [eds.], *The timetree of life*. Oxford University Press, New York. 551 p.
- HERBERT, J., M. W. CHASE, M. MÖLLER, AND R. J. ABBOTT. 2006. Nuclear and plastid DNA sequences confirm the placement of the enigmatic *Canacomyrica monticola* in Myricaceae. *Taxon* **55**: 349–357.
- IVANCIUC, O., C. H. SCHEIN, AND W. BRAUN. 2003. SDAP: database and computational tools for allergenic proteins. *Nucl. Acids Res.* **31**: 359–362.
- JENKINS, J. A., S. GRIFFITHS-JONES, P. R. SHEWRY, H. BREITENEDER, AND E. N. C. MILLS. 2005. Structural relatedness of plant food allergens with specific reference to cross-reactive allergens: an in silico analysis. *Journal of Allergy and Clinical Immunology* **115**: 163–170.
- KLEBER-JANKE, T., R. CRAMERI, U. APPENZELLER, M. SCHLAACK, AND W.-M. BECKER. 1999. Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology. *International Archives of Allergy and Immunology* **119**: 265–274.
- KRAUSE, S., G. REESE, S. RANDOW, D. ZENNARO, D. QUARATINO, P. PALAZZO, M. A. CIARDIELLO, A. PETERSEN, W.-M. BECKER, AND A. MARI. 2009. Lipid transfer protein (Ara h 9) as a new peanut allergen relevant for a Mediterranean allergic population. *Journal of Allergy and Clinical Immunology* **124**: 771–778.

- KULIS, M., C. K. WAN, B. K. GORENTLA, A. W. BURKS, AND X. P. ZHONG. 2011. Diacylglycerol kinase zeta deficiency in a non-CD4+ T cell compartment leads to increased peanut hypersensitivity. *Journal of Allergy and Clinical Immunology* **128**: 212.
- LATCHAM, F., L. MERINO, A. LANG, J. GARVEY, M. A. THOMSON, J. A. WALKER-SMITH, S. E. DAVIES, A. D. PHILLIPS, AND S. H. MURCH. 2003. A consistent pattern of minor immunodeficiency and subtle enteropathy in children with multiple food allergy. *Journal of Pediatrics* **143**: 39–47.
- LEHRER, S. B., R. AYUSO, AND G. REESE. 2006. Current understanding of food allergens. *Ann. New York Acad. Sci.* **964**: 69–85.
- LEHTO, M., L. AIRAKSINEN, A. PUUSTINEN, S. TILLANDER, S. HANNULA, T. NYMAN, E. TOSKALA, H. ALENUS, AND A. LAURMA. 2010. Thaumatin-like protein and baker's respiratory allergy. *Annals of Allergy, Asthma & Immunology* **104**: 139–146.
- LEPOITTEVIN, J. P., C. BENEZRA, AND Y. ASAKAWA. 1989. Allergic contact dermatitis to *Ginkgo biloba* L.: relationship with urushiol. *Archives of Dermatological Research* **281**: 227–230.
- LERCH, M., C. EGGER, AND A. J. BIRCHER. 2005. Allergic reactions to macadamia nut. *Allergy* **60**: 130–131.
- LI, R. Q., Z. D. CHEN, A. M. LU, D. E. SOLTIS, P. S. SOLTIS, AND P. MANOS. 2004. Phylogenetic relationships in Fagales based on DNA sequences from three genomes. *Int. J. Pl. Sci.* **165**: 311–324.
- MAGALLON, S., K. W. HILU, AND D. QUANDT. 2013. Land plant evolutionary timeline: gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *Amer. J. Bot.* **100**: 556–573.
- MANOS, P. S. AND K. P. STEELE. 1997. Phylogenetic analyses of "higher" Hamamelididae based on plastid sequence data. *Amer. J. Bot.* **84**: 1407–1419.
- MARINAS, D., L. VILA, AND M. L. SANZ. 1998. Allergy to pine nuts. *Allergy* **53**: 220–222.
- MATHEWS, S. 2009. Phylogenetic relationships among seed plants: persistent questions and the limits of molecular data. *Amer. J. Bot.* **96**: 228–236.
- MILLER, M. A., W. PFEIFFER, AND T. SCHWARTZ. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (CGE), 14 Nov 2010, New Orleans, LA, pp. 1–8.
- MILLS, E. N. C., M. J. C. JENKINS, P. R. ALCOCER, AND P. R. SHEWRY. 2004. Structural, biological, and evolutionary relationships of plant food allergens sensitizing via the gastrointestinal tract. *Critical Reviews in Food Science and Nutrition* **44**: 379–407.
- MITTAG, D., J. AKKERDAAS, B. K. BALLMER-WEBER, L. VOGEL, M. WENSING, W.-M. BECKER, S. J. KOPPELMAN, A. C. KNULST, A. HELBLING, S. L. HEFLE, R. VAN REE, AND S. VIETHS. 2004. Ara h 8, a Bet v 1-homologous allergen from peanut, is a major allergen in patients with combined birch pollen and peanut allergy. *Journal of Allergy and Clinical Immunology* **114**: 1410–1417.
- MONERET-VAUTRIN, D. A., F. RANCE, G. KANNY, A. OLSEWSKI, J. L. GUEANT, G. DUTAU, AND L. GUERIN. 1998. Food allergy to peanuts in France—evaluation of 142 observations. *Clinical & Experimental Allergy* **28**: 1113–1119.
- MORAFO, V., K. SRIVASTAVA, C. K. HUANG, G. KLEINER, S. Y. LEE, H. A. SAMPSON, AND X. M. LI. 2003. Genetic susceptibility to food allergy is linked to differential TH2-TH1 responses in C3H/HeJ and BALB/c mice. *Journal of Allergy and Clinical Immunology* **111**: 1122–1128.
- MORGAN, W. A. AND B. J. CLAYSHULTE. 2000. Pecans lower low-density lipoprotein cholesterol in people with normal lipid levels. *Journal of the American Dietetic Association* **100**: 312–318.
- NGUYEN, S. A., D. R. MORE, B. A. WHISMAN, AND L. L. HAGAN. 2004. Cross-reactivity between coconut and hazelnut proteins in a patient with coconut anaphylaxis. *Annals of Allergy, Asthma & Immunology* **92**: 281–284.
- NIELSEN, N. H. 1990. Systemic allergic reaction to pine nuts. *Ann. Allergy* **64**: 132–133.
- NOYES, J. N., G. K. BOYD, AND G. A. SETTIPANE. 1979. Anaphylaxis to sunflower seed. *Journal of Allergy and Clinical Immunology* **63**: 242–244.
- OSBORNE, N. J., J. J. KOPLIN, P. E. MARTIN, L. C. GURRIN, A. J. LOWE, M. C. MATHESON, A.-L. PONSONBY, M. WAKE, M. L. K. TANG, S. C. DHARMAGE, AND K. J. ALLEN. 2011. Prevalence of challenge-proven IgE-mediated food allergy using population-based sampling and predetermined challenge criteria in infants. *Journal of Allergy and Clinical Immunology* **127**: 668–676.
- PASTORELLO, E. A., E. VARIN, L. FARIOLI, V. PRAVETTONI, C. ORTOLANI, C. TRAMBAIOLI, D. FORTUNATO, M. G. GIUFFRIDA, F. RIVOLTA, A. ROBINO, A. M. CALAMARI, L. LACAVA, AND A. CONTI. 2001. The major allergen of sesame seeds (*Sesamum indicum*) is a 2S albumin. *Journal of Chromatography B: Biomedical Sciences and Applications* **756**: 85–93.
- POLTRONIERI, P., M. S. CAPPELLO, N. DOHMAE, A. CONTI, D. FORTUNATO, E. A. PASTORELLO, C. ORTOLANI, AND G. ZACHEO. 2002. Identification and characterisation of the IgE-binding proteins 2S albumin and conglutin gamma in almond (*Prunus dulcis*) seeds. *International Archives of Allergy and Immunology* **128**: 97–104.
- RABJOHN, P., E. M. HELM, J. S. STANLEY, C. M. WEST, H. A. SAMPSON, A. W. BURKS, AND G. A. BANNON. 1999. Molecular cloning and epitope analysis of the peanut allergen Ara h 3. *Journal of Clinical Investigation* **103**: 535–542.
- RADAUER, C. AND H. BREITENEDER. 2007. Evolutionary biology of plant food allergens. *Journal of Allergy and Clinical Immunology* **120**: 518–525.
- , M. BUBLIN, S. WAGNER, A. MARI, AND H. BREITENEDER. 2008. Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *Journal of Allergy and Clinical Immunology* **121**: 847–852.
- , M. WILLERROIDER, H. FUCHS, K. HOFFMANN-SOMMERGRUBER, J. THALHAMER, F. FERREIRA, O. SCHEINER, AND H. BREITENEDER. 2006. Cross-reactive and species-specific immunoglobulin E epitopes of plant profilins: an experimental and structure-based analysis. *Clinical & Experimental Allergy* **36**: 920–929.
- RAN, J.-H., H. GAO, AND X.-Q. WANG. 2010. Fast evolution of the retroprocessed mitochondrial rps3 gene in Conifer II and further evidence for the phylogeny of gymnosperms. *Molec. Phylogen. Evol.* **54**: 136–149.
- RIHS, H.-P., P. ROZYNEK, K. MAY-TAUBE, B. WELTICKE, AND X. BAUR. 1994. Polymerase chain reaction based cDNA cloning of wheat profilin: a potential plant allergen. *International Archives of Allergy and Immunology* **105**: 190–194.
- RIHS, H.-P., Z. CHEN, F. RUÉFF, A. PETERSEN, P. ROZYNEK, H. HEIMANN, AND X. BAUR. 1999. IgE binding of the recombinant allergen soybean profilin (rGly m 3) is mediated by conformational epitopes. *Journal of Allergy and Clinical Immunology* **104**: 1293–1301.
- ROBOTHAM, J. M., F. WANG, V. SEAMON, S. S. TEUBER, S. K. SATHE, H. A. SAMPSON, K. BEYER, M. SEAVY, AND K. H. ROUX. 2005. Ana o 3, an important cashew nut (*Anacardium occidentale* L.) allergen of the 2S albumin family. *Journal of Allergy and Clinical Immunology* **115**: 1284–1290.
- , S. S. TEUBER, S. K. SATHE, AND K. H. ROUX. 2002. Linear IgE epitope mapping of the English walnut (*Juglans regia*) major food allergen, Jug r 1. *Journal of Allergy and Clinical Immunology* **109**: 143–149.
- ROSENGARTEN, F. 1984. The book of edible nuts. Walker & Co, New York.
- ROUX, K. H., S. S. TEUBER, AND S. K. SATHE. 2003. Tree nut allergens. *International Archives of Allergy and Immunology* **131**: 234–244.

- ROUX, N., S. HOGENDIJK, AND C. HAUSER. 1998. Severe anaphylaxis to pine nuts. *Allergy* **53**: 213–222.
- RUIQI, L., C. ZHIDUAN, H. YAPING, AND L. ANMING. 2002. Phylogenetic relationships of the “higher” hamamelids based on chloroplast trnL-F sequences. *Acta Bot. Sin.* **44**: 1462–1468.
- SALCEDO, G., A. DIAZ-PERALES, AND R. SANCHEZ-MONGE. 2001. The role of plant panallergens in sensitization to natural rubber latex. *Current Opinion in Allergy and Clinical Immunology* **1**: 177–183.
- SANDER, I., P. ROZYNEK, H.-P. RIHS, V. VAN KAMPEN, F. T. CHEW, W. S. LEE, N. KOTSCHY-LANG, R. MERGET, T. BRÜNING, AND M. RAULF-HEIMSOOTH. 2011. Multiple wheat flour allergens and cross-reactive carbohydrate determinants bind IgE in baker’s asthma. *Allergy* **66**: 1208–1215.
- SAUQUET, H., S. Y. HO, M. A. GANDOLFO, G. J. JORDAN, P. WILF, D. J. CANTRILL, M. J. BAYLY, L. BROMHAM, G. K. BROWN, R. J. CARPENTER, D. M. LEE, D. J. MURPHY, J. M. K. SNIDERMAN, AND F. UDOVICIC. 2012. Testing the impact of calibration on molecular divergence times using a fossil-rich group: the case of *Nothofagus* (Fagales). *Syst. Biol.* **61**: 289–313.
- SDAP 2015. Structural Database of Allergenic Proteins. <http://fermi.utmb.edu> (Mar 2015).
- SHARMA, G. M., A. IRSIGLER, P. DHANARAJAN, R. AYUSO, L. BARDINA, H. A. SAMPSON, K. H. ROUX, AND S. K. SATHE. 2011. Cloning and characterization of an 11S Legumin, Car i 4, a major allergen in pecan. *J. Agric. Food Chem.* **59**: 9542–9552.
- SHIN, D. S., C. M. COMPADRE, S. J. MALEKI, R. A. KOPPER, H. SAMPSON, S. K. HUANG, A. W. BURKS, AND G. A. BANNON. 1998. Biochemical and structural analysis of the IgE binding sites on Ara h1, an abundant and highly allergenic peanut protein. *J. Biol. Chem.* **273**: 13753–13759.
- SICHERER, S. 2002. Clinical update on peanut allergy. *Annals of Allergy, Asthma & Immunology* **88**: 350–361.
- , A. W. BURKS, AND H. A. SAMPSON. 1998. Clinical features of acute allergic reactions to peanut and tree nuts in children. *Pediatrics* **102**: 1–6.
- SOLTIS, D. E., P. S. SOLTIS, M. E. CHASE, M. E. MORT, D. C. ALBACH, M. ZANIS, V. SAVOLAINEN, W. H. HAHN, S. B. HOOT, M. F. FAY, M. AXTELL, S. M. SWENSEN, L. M. PRINCE, W. J. KRESS, K. C. NIXON, AND J. S. FARRIS. 2000. Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. *Bot. J. Linn. Soc.* **133**: 381–461.
- STAMATAKIS, A., A. J. ABERER, C. GOLL, S. A. SMITH, S. A. BERGER, AND F. IZQUIERDO-CARRASCO. 2012. RAXML-Light: a tool for computing terabyte phylogenies. *Bioinformatics* **28**: 2064–2066.
- STEIN, W. E., C. M. BERRY, L. V. HERNICK, AND F. MANNOLINI. 2012. Surprisingly complex community discovered in the mid-Devonian fossil forest at Gilboa. *Nature* **483**: 78–81.
- STEVENS, P. 2001 onwards. Angiosperm Phylogeny Website, vers. 12 [and continuously updated since]. <http://www.mobot.org/MOBOT/research/APweb/welcome.html> (Mar 2015).
- STÖVER, B. C. AND K. F. MÜLLER. 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics* **11**: 7.
- STUTIUS, L. M., W. J. SHEEHAN, P. RANGSITHIENCHAI, A. BHARMANEE, J. E. SCOTT, M. C. YOUNG, A. F. DIJON, L. C. SCHNEIDER, AND W. PHIPATANAKUL. 2010. Characterizing the relationship between sesame, coconut, and nut allergy in children. *Pediatric Allergy and Immunology* **21**: 1114–1118.
- SUTHERLAND, M. F., R. E. O’HEHIR, D. CZARNY, AND C. SUPHIOGLU. 1999. Macadamia nut anaphylaxis: demonstration of specific IgE reactivity and partial cross-reactivity with hazelnut. *Journal of Allergy and Clinical Immunology* **104**: 889–890.
- TARIQ, S. M., M. STEVENS, S. MATTHEWS, S. RIDOUT, R. TWISELTON, AND D. W. HIDE. 1996. Cohort study of peanut and tree nut sensitisation by age of 4 years. *BMJ* **313**: 514–517.
- TEUBER, S. S. AND W. R. PETERSON. 1999. Systemic allergic reaction to coconut (*Cocos nucifera*) in 2 subjects with hypersensitivity to tree nut and demonstration of cross-reactivity to legumin-like seed storage proteins: new coconut and walnut food allergens. *Journal of Allergy and Clinical Immunology* **103**: 1180–1185.
- , S. S. COMSTOCK, S. K. SATHE, AND K. H. ROUX. 2003. Tree nut allergy. *Current Allergy and Asthma Reports* **3**: 54–61.
- , A. M. DANDEKAR, W. R. PETERSON, AND C. L. SELLERS. 1998. Cloning and sequencing of a gene encoding a 2S albumin seed-storage protein precursor from English walnut (*Juglans regia*), a major food allergen. *Journal of Allergy and Clinical Immunology* **101**: 807–814.
- , K. C. JARVIS, A. M. DANDEKAR, W. R. PETERSON, AND A. A. ANSARI. 1999. Identification and cloning of a complementary DNA encoding a vicilin-like proprotein, jug r 2, from English walnut kernel (*Juglans regia*), a major food allergen. *Journal of Allergy and Clinical Immunology* **104**: 1311–1320.
- USDA [United States Department of Agriculture]. 2014. http://www.nass.usda.gov/Statistics_by_State/California/Publications/California_Ag_Statistics/Reports/ (Mar 2015).
- VIETHS, S., S. SCHEURER, AND B. K. BALLMER-WEBER. 2002. Current understanding of cross-reactivity of food allergens and pollen. *Ann. New York Acad. Sci.* **964**: 47–68.
- WALLOWITZ, M., W. R. PETERSON, S. URATSU, S. S. COMSTOCK, A. M. DANDEKAR, AND S. S. TEUBER. 2006. Jug r 4, a legumin group food allergen from walnut (*Juglans regia* cv. Chandler). *J. Agric. Food Chem.* **54**: 8369–8375.
- WANG, F., J. M. ROBOTHAM, S. S. TEUBER, P. TAWDE, S. K. SATHE, AND K. H. ROUX. 2002. Ana o 1, a cashew (*Anacardium occidentale*) allergen of the vicilin seed storage protein family. *Journal of Allergy and Clinical Immunology* **110**: 160–166.
- WILLISON, L. N., P. TAWDE, J. M. ROBOTHAM, R. M. PENNEY, IV, S. S. TEUBER, S. K. SATHE, AND K. H. ROUX. 2008. Pistachio vicilin, Pis v 3, is immunoglobulin E-reactive and cross-reacts with the homologous cashew allergen, Ana o 1. *Clinical & Experimental Allergy* **38**: 1229–1238.
- , P. TRIPATHI, G. M. SHARMA, S. S. TEUBER, S. K. SATHE, AND K. H. ROUX. 2011. Cloning, expression and patient IgE reactivity of recombinant Pru du 6, an 11S globulin from almond. *International Archives of Allergy and Immunology* **156**: 267–281.
- WITKE, W. 2004. The role of profilin complexes in cell motility and other cellular processes. *Trends Cell Biol.* **14**: 461–469.
- YAGAMI, A. 2010. Anaphylaxis to lipid transfer protein from sunflower seeds. *Allergy* **65**: 1340–1341.
- YANG, Z.-Y., J.-H. RAN, AND X.-Q. WANG. 2012. Three genome-based phylogeny of Cupressaceae s.l.: further evidence for the evolution of gymnosperms and Southern Hemisphere biogeography. *Molec. Phylog. Evol.* **64**: 452–470.
- YUNGINGER, J. W. 1988. Fatal food-induced anaphylaxis. *JAMA* **260**: 1450.
- ZUIDMEER, L., K. GOLDBAHN, R. J. RONA, D. GISLASON, C. MADSEN, C. SUMMERS, E. SODERGREN, J. DAHLSTROM, T. LINDNER, S. T. SIGURDARDOTTIR, D. MCBRIDE, AND T. KEIL. 2008. The prevalence of plant food allergies: a systematic review. *Journal of Allergy and Clinical Immunology* **121**: 1210–1218.

APPENDIX 1

GLOSSARY

- Allergenic protein:** A protein that induces an allergic response in a patient.
- Allergic reaction:** An abnormal immune response induced by exposure to an allergen (antigen).

Cross-reactivity: When an IgE antibody can recognize similar allergens from different species that share the same epitopes. Tested in vitro with an IgE blood test.

Cross-sensitivity: Sensitivity to more than one allergen, usually due to the display of similar epitopes, which provokes cross reactivity between the two allergens. Can be assayed with a skin prick test.

Co-allergy: Allergic reaction to more than one species that is due to separate IgE antibody responses.

Epitope: The specific region of the protein's amino acid sequence (the antigen) that is recognized by the immune system.

IgE: Immunoglobulin E; a subtype of antibody that plays a major role in allergic diseases.

Nut: A botanical term that describes a specific type of fruit that contains a single seed, with a hard, dry, outer layer, and a special covering called a cupule (the "cap" of an acorn).

Serum: The clear liquid that separates from blood when it is allowed to clot completely, used in blood allergy testing because it contains a patient's IgE.