



Spontaneous and induced variability of allergens in commodity crops: Ara h 2 in peanut as a case study

Peggy Ozias-Akins^{a,*}, M. Laura Ramos^a, Paola Faustinelli^a, Ye Chu^a, Soheila Maleki^b, Jay J. Thelen^c, James Huntley^d, Katherine Arias^e, Manel Jordana^e

^a Department of Horticulture, University of Georgia Tifton Campus, 2356 Rainwater Road, Tifton, GA 31793-0748, USA

^b USDA-ARS-SRRC, New Orleans, LA 70124, USA

^c Department of Biochemistry, University of Missouri-Columbia, Columbia, MO 65211, USA

^d National Center for Genome Resources, Santa Fe, NM 87505, USA

^e Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ont., Canada

ARTICLE INFO

Article history:

Received 5 September 2008

Available online 20 November 2008

Keywords:

Peanut

Groundnut

Arachis duranensis

Arachis hypogaea

Ara h 2

RNA interference

RNA silencing

TILLING

Mutagenesis

ABSTRACT

Many commodity crops are grown for human consumption, and the resulting food products usually contain proteins, some of which may be allergenic. The legumes, peanut and soybean, as well as tree nuts and some cereal grains are well recognized sources of food allergens. In peanut, there are 11 documented allergenic proteins, although the major allergens are considered to be Ara h 1 and Ara h 2, both of which are seed storage proteins. Methods to reduce or eliminate these proteins from seeds are available and allow the feasibility of this approach to be tested. Greatly reduced amounts of Ara h 2 can be achieved by RNA silencing in transgenic peanut; however, mutagenesis is a more viable and socially acceptable approach to allergen elimination. Although the techniques for mutagenesis are not new, methods for mutant detection at the molecular level have recently been developed. However, these methods are dependent on genome sequence. These methods will facilitate discovery of spontaneous and induced mutations that may be useful over the long term to eliminate certain allergens from peanut.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Crops with significant food allergen issues include the legumes, peanut and soybean, cereal grains such as wheat, and tree nuts (walnut, Brazil nut, among other phylogenetically diverse species) (Teuber et al., 2006). Officially recognized allergenic proteins may include one or multiple proteins per species (www.allergen.org; www.allergome.org) and fall into a variety of functional/biochemical classes (Ozias-Akins et al., 2006). These classes include seed storage proteins (e.g., globulins, albumins), metabolic proteins such as profilin (an actin binding protein), and certain proteins with protective functions such as insect or fungal resistance (agglutinin, protease inhibitors, lipid transfer proteins, etc.). Natural genetic variation can produce multiple isoforms for one allergen, and these isoforms can be officially recognized as distinct if sufficient evidence for allergenicity is provided to the Allergen Nomenclature Sub-Committee of the International Union of Immunological Societies (I.U.I.S.). The committee considers submissions that meet the following criteria: reactivity with IgE of more than

5% prevalence and a minimum of five patients that show IgE reactivity. We are exploring molecular variation for allergens in peanut (*Arachis hypogaea* L.) with the goal of using genetic resources to reduce allergenicity of the food product.

2. Peanut allergens and their genes

Peanut seeds contain 11 officially recognized allergens, Ara h 1–Ara h 11. Ara h 1 and Ara h 2 are considered to be the most clinically relevant allergens because >90% of peanut allergic individuals have IgE that reacts with these proteins (Burks et al., 1998), although Ara h 3 is a significant allergen among some population subgroups (Restani et al., 2005). Allergens are named according to genus and species of the source, in the order they were found. For example, the second discovered allergen from peanut (*Arachis hypogaea*) was named **Ara h 2**. This allergenic protein was first described as two protein bands of approximately 17 kDa according to polyacrylamide gel electrophoresis (Burks et al., 1992), and structural analysis indicated that eight cysteine residues allow the formation of four disulfide bonds that contribute to protease resistance (Sen et al., 2002). It was later shown that the two bands observed on polyacrylamide gels actually represented two iso-

* Corresponding author. Fax: +229 386 7371.

E-mail address: pozias@uga.edu (P. Ozias-Akins).

forms that differed most significantly by a 12 amino acid insertion/deletion in the coding region (Chatel et al., 2003). Ramos et al. (2006) assigned the two isoforms to the two subgenomes of tetraploid peanut (Fig. 1) by studying their similarity with homologous sequences in the diploid progenitors of peanut, *Arachis duranensis* (A genome) and *Arachis ipaensis* (B genome). The lower molecular weight fragment corresponds to the A genome isoform. Nucleotide and predicted amino acid sequences for Ara d 2.01 [Accession Nos. EF609641 (nucleotide) and ABQ96212 (protein)] and Ara i 2.02 [Accession Nos. EF609642 (nucleotide) and ABQ96213 (protein)] show that the wild diploid species have forms of the genes that are essentially identical to their orthologs in *A. hypogaea*. Their allergenic properties, therefore, should be the same.

3. Screening for spontaneous and induced mutational variation in allergen genes

The putative diploid progenitors of peanut that we initially studied, *A. duranensis* and *A. ipaensis*, were the same accessions identified by Kochert et al. (1996) as being most similar to cultivated tetraploid peanut, *A. hypogaea*, for molecular marker polymorphisms. Our observation that orthologous *Ara h 2* sequences from the progenitor genomes are identical to cultivated peanut sequences further supports their conclusions. However, they also observed considerable polymorphism among different accessions of *A. duranensis*. In order to search for polymorphisms in the A genome, *Ara d 2.01* gene (*Ara h 2.01* ortholog), we screened 30 accessions of *A. duranensis* using a technique called EcoTILLING (Comai et al., 2004), which is based on the TILLING (targeting induced local lesions in genomes) method for detecting induced mutations (Henikoff and Comai, 2003). Briefly, DNA is extracted from individ-

uals that potentially contain induced or spontaneous mutations and pooled 2-, 4- or 8-fold. Each pool should then contain only wild-type or wild-type plus mutant alleles of a gene of interest. The gene sequence information must be available to allow PCR primer design such that a 0.5–1.5 kb amplicon can be generated using infrared dye labeled primers. Labeled amplicons from pooled DNAs will contain a mutation if present in the pool, and mutant plus wild-type amplicons can reanneal as heteroduplexes. Single nucleotide mismatches are detected and cleaved by endonucleases such as the enzyme Cel I, and fragments are separated by denaturing polyacrylamide gel electrophoresis, usually on a Licor DNA analysis system where both fluorescent labels can be detected. Differentially labeled fragments that represent the same mutation should be complementary in size. Once a pool of DNAs is identified to contain a putative mutant allele, the process is repeated with individual DNAs to verify which one is the mutant. The mutant allele DNA sequence then is determined by routine sequencing after which the impact of the mutation on amino acid sequence can be predicted. Ideally a collection of different alleles can be obtained and used to study gene function.

After EcoTILLING of *A. duranensis* for *Ara d 2.01* mutations and sequencing of mutant alleles, seven accessions displayed altogether eight single nucleotide polymorphisms (SNPs) in the coding region, five of which resulted in amino acid substitutions (Ramos et al., 2008). Three of these substitutions were in previously identified IgE epitopes (Stanley et al., 1997). Although none was predicted to radically alter tertiary structure, one change was found to result in significantly reduced IgE binding from three patients tested. This most interesting mutation was present in two *A. duranensis* accessions, both originating from Paraguay. The reduction in IgE binding quantified using densitometry of signals from replicated samples on western blots ranged from 56% to 99%. The question remains whether this orthologous isoform (or natural variant) of *Ara h 2* would behave differently from wild-type for sensitization and allergenicity in a mouse or swine model, in basophil activation assays, or in human skin-prick tests. If a less allergenic isoform were to be identified in peanut relatives, it would be feasible, although long-term, to introgress the mutant allele into cultivated peanut through interspecific hybridization (Simpson 2001). An alternative to introgression might be to induce mutation to a similar allele by chemical mutagenesis or ionizing radiation of cultivated peanut (Ozias-Akins et al., 2006).

TILLING also has been used to screen for induced mutations in cultivated peanut. Peanut seeds are treated with ethylmethanesulfonate or similar DNA alkylating agents that cause G/C to A/T transitions (Greene et al., 2003) and planted to generate the M1 generation. A single M2 seed is harvested from each plant and DNA is extracted from leaves of M2 plants. TILLING is conducted on M2 leaf DNA and M3 seeds from each plant are stored for later sampling after mutant lines have been identified. The efficiency of the screening largely depends on pooling depth. Populations that are of manageable sizes to screen for mutants in multiple genes should have a mutation frequency of 10^{-5} or greater (Henikoff et al., 2004). A small mutant population of cultivated peanut has allowed the identification of three mutations in *Ara h 2.01* neither of which is a knockout mutation, but two of them cause a predicted amino acid change. Screening in cultivated peanut was designed to allow *Ara h 2.01* and *2.02* to be separately screened using gene-specific primers. Since *Ara h 2* genes are small and without introns, primers were targeted to non-coding upstream and downstream regions, allowing amplification of 1.2 kb encompassing the entire coding region. In order to pursue this strategy for peanut on a large scale, much more genomic sequence information will be required. At this time limited sequence information is available and is largely restricted to expressed sequence tag (EST) sequences from seed and leaf tissues.

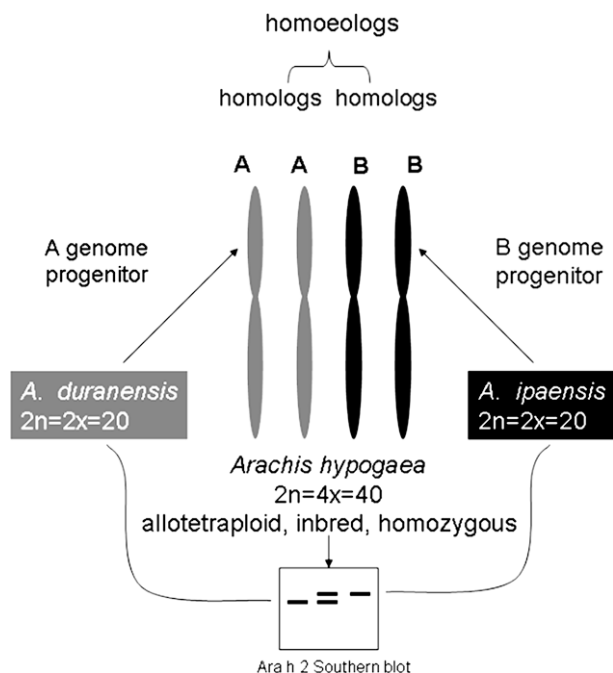


Fig. 1. Representative chromosomes from each of the four sets of 10 chromosomes found in allotetraploid, cultivated peanut. The paired chromosomes are homologs from the A and B subgenomes, and the A and B genome chromosomes are homeologous. *Ara h 2* has been shown to be encoded by a single gene in each genome (Ramos et al., 2006). The two *Ara h 2* alleles on the A genome homologs are identical to one another (homozygous) as are those on the B genome homologs; however, the A and B genome *Ara h 2* genes in cv. Georgia Green differ from each other by a 36 bp insertion/deletion and eight single nucleotide polymorphisms. The 36 bp indel results in an additional 12 amino acids in *Ara h 2.02* compared with *Ara h 2.01* and two of the SNPs result in conservative amino acid substitutions.

While TILLING for *Ara h 2* mutants is relatively straightforward because of a single gene copy in each of the subgenomes of cultivated peanut and *Ara h 1* shows a similar level of complexity (unpublished results), *Ara h 3* is likely to be encoded by multiple members within a gene family since more than two genes already have been estimated from Southern blots (Kang and Gallo, 2007). Even for *Ara h 2*, there is a very low probability that a mutation in each of the two genes could be found in a single mutant individual; therefore, it will be necessary to combine two mutations by conventional breeding and marker-assisted selection. The scenario becomes more complex for genes with even more copies. Hence TILLING for a reduced allergen peanut is a long-term project but probably the most effective means to ensure stable allergen reduction.

4. Variation resulting from allergen gene silencing

The clinical consequences of allergen reduction can be studied in the shorter term by applying gene silencing methodology. Post-transcriptional gene silencing is now known to be the consequence of an evolutionarily ancient natural RNA surveillance system. Foreign or aberrant transcripts can be recognized and along with their homologous sequences can be targeted for degradation (primarily in plants) or translational inhibition (primarily in animals) (Vaucheret 2006). Since we know the sequence of *Ara h 2* transcripts, we have used sequence-dependent gene silencing to knockdown expression of *Ara h 2* in peanut seeds. This was accomplished by transforming peanut with a silencing construct containing a portion of the *Ara h 2* gene in inverted repeat orientation and separated by an intervening sequence (Chu et al., 2008). Such an inverted repeat construct design allows RNA to form a stem-loop structure that is recognized as aberrant and “diced” into ~24 nt fragments. These short fragments will anneal with homologous RNAs, which also become targeted for degradation. Silencing of gene (and by extension protein) expression can be fractional or nearly complete. Certain viruses have evolved mechanisms to counter silencing with suppressor proteins (Roth et al., 2004). Because of this possible mechanism of instability, among others associated with transgene instability, silencing of major food allergen genes via RNAi is not considered to be a commercially viable solution. Nevertheless, RNAi is an extremely valuable approach to test the effect of specific protein reduction on peanut seed development and viability as well as for allergenicity studies using animal models. We have begun a systematic approach to test the effects of reducing peanut seed allergens individually and in combination.

Peanut transformation can be routinely accomplished by micro-projectile bombardment of embryogenic tissue cultures (Ozias-Akins 2007). Transformation efficiency is low compared to many other crops but sufficient to allow the recovery of multiple independent transgenic events. From our work to silence *Ara h 2*, three events have been verified as independent based on Southern blot analysis and all three show a certain degree of *Ara h 2* protein reduction (Fig. 2). As expected, IgE from human serum of peanut allergic individuals also detected the reduction in *Ara h 2* protein. Based on 1-D and 2-D difference gel electrophoresis patterns and IgE binding on Western blots, there were no apparent major changes in other seed protein patterns (Chu et al., 2008). Two samples of defatted crude peanut extract, one with low and one with wild-type levels of *Ara h 2* were tested in a murine model for sensitization capacity and allergic response. No significant difference was observed in allergic reaction of mice either: (1) sensitized and challenged with the wild-type *Ara h 2* sample, or (2) sensitized with low-level *Ara h 2* and challenged with low or wild-type samples. Our conclusions based on the current data are that: (1) the low quantitative threshold for allergic response to *Ara h 2* may

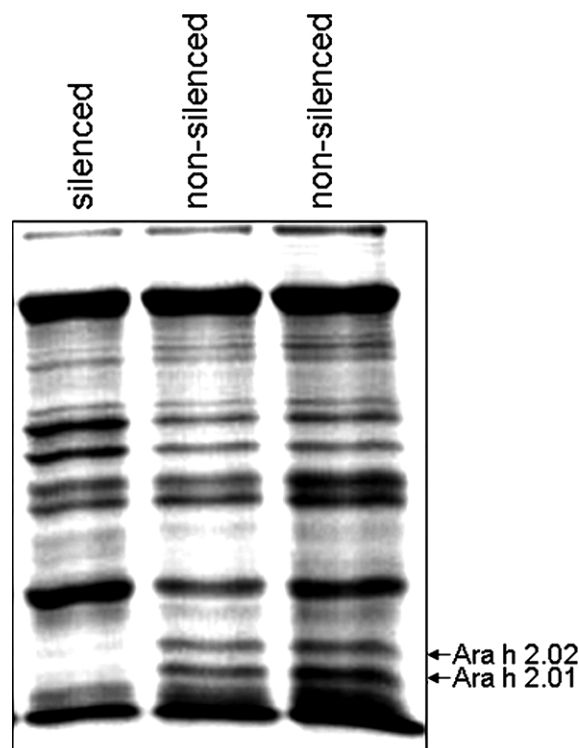


Fig. 2. Polyacrylamide gel electrophoresis (1-D) of total protein extracted from peanut seeds. Each lane represents extract from a single seed with all seeds derived from a single transgenic event. The parent plant was heterozygous (hemizygous) for the transgene; therefore, the trait for *Ara h 2* silencing is segregating among progeny. The lane labeled “silenced” has significantly lower amounts of both *Ara h 2* isoforms than “non-silenced”.

not have been reached with the silenced line tested, or (2) the presence of wild-type levels of *Ara h 1*, *Ara h 6*, and perhaps *Ara h 3* may mask any effect of reduced *Ara h 2* in the mouse model. The human biological relevance of murine models to test for allergenicity is debatable; therefore, additional types of analyses should be conducted before strong conclusions are drawn. For example, skin prick tests (SPT) on subjects monosensitized to *Ara h 2* (Astier et al., 2006), or measurement of mediator release after challenge of humanized rat basophilic leukemia (RBL) cells (Lehmann et al., 2006) could better predict whether peanut with reduced *Ara h 2* would be hypoallergenic.

5. Significance of genetic variability

There are several questions that the variability we have found or generated in peanut can help to address, particularly if animal models are shown to be valid surrogates for human allergenicity testing. These are related to the extent that amino acid substitutions in allergen proteins can affect their properties, either protein stability, function or allergenicity/reactivity. We have shown that a single amino acid change naturally occurring in an accession of the A genome progenitor of peanut significantly reduces IgE binding to the linear epitope. This change occurs in the immunodominant epitope #7. Previous mutational analysis of *Ara h 2* epitopes actually had shown an increase in IgE binding with a different amino acid substitution at this site (Stanley et al., 1997). The secondary and tertiary protein structure most likely is unaffected based on evidence from computational modeling and no apparent change in stability of the protein in vivo. Another example of a natural genetic variant that shows an even more dramatic consequence for protein expression is from soybean. The allergenic protein P34, a

cysteine protease, was absent from two out of 14,239 cultivated soybean (*Glycine max*) and seven out of 1114 wild relative (*Glycine soja*) accessions screened (Joseph et al., 2006). The two soybean accessions had P34 genes of identical sequence that displayed six mutations when compared with the wild-type sequence. Four of the mutations were predicted to result in amino acid substitutions and one of the substitutions was from serine to cysteine. Since a full-length transcript of the variant gene was detected on northern blots but no protein was produced, it was speculated that the additional cysteine probably disrupted tertiary structure and caused protein instability. For peanut, TILLING of induced mutants also has the potential to identify an amino acid substitution that would add or remove a cysteine from Ara h 2 and alter its stability by changing disulfide bonding. Although we can speculate on the consequences of these variant genes/proteins for allergenicity of the food product, testing of the food or total proteins from the food in animal models or by using SPT or RBL assays would be necessary steps to obtain definitive data.

The development of various animal models is another topic of discussion at this workshop. Such testing should also take into consideration the variation in food processing methods used for a commodity such as peanut. For example, peanuts can be eaten raw, roasted, boiled or processed into a variety of products. Roasting is commonly used during processing of peanuts in the US and roasting has been shown to enhance binding of IgE to peanut proteins, probably by promoting the formation of products of the Maillard reaction between sugar and protein (Maleki et al., 2000; Chung et al., 2003; Gruber et al., 2005). Can genetic variants respond differently to roasting either in terms of interactions with other macromolecules or downstream digestion? This and related questions can be addressed by the abundant research tools now available to facilitate discovery of genetic variation (DNA sequence and analysis tools), its consequence for protein variation (proteomics tools) and its consequence for allergenicity (animal models).

References

- Astier, C., Morisset, M., Roitel, O., Codreanu, F., Jacquenet, S., Franck, P., Ogier, V., Petit, N., Proust, B., Moneret-Vautrin, D., Burks, A.W., Bihain, B., Sampson, H.A., Kanny, G., 2006. Predictive value of skin prick tests using recombinant allergens for diagnosis of peanut allergy. *J. Allergy Clin. Immunol.* 118, 250–256.
- Burks, A.W., Williams, L.W., Connaughton, C., Cockrell, G., O'Brien, T.J., Helm, R.M., 1992. Identification and characterization of a second major peanut allergen, Ara h II, with use of the sera of patients with atopic dermatitis and positive peanut challenge. *J. Allergy Clin. Immunol.* 90, 962–969.
- Burks, W., Sampson, H.A., Bannon, G.A., 1998. Peanut allergens. *Allergy* 53, 725–730.
- Chatel, J.M., Bernard, H., Orson, F.M., 2003. Isolation and characterization of two complete Ara h 2 isoforms cDNA. *Int. Arch. Allergy Immunol.* 131, 14–18.
- Chu, Y., Faustinelli, P., Ramos, M.L., Hajdich, M., Stevenson, S., Thelen, J.J., Maleki, S.J., Ozias-Akins, P., 2008. Reduction of IgE binding and non-promotion of *Aspergillus flavus* fungal growth by simultaneously silencing Ara h 2 and Ara h 6 in peanut. *J. Agric. Food Chem.*, doi:10.1021/jf802600r.
- Chung, S.-Y., Butts, C.L., Maleki, S.J., Champagne, E.T., 2003. Linking peanut allergenicity to the processes of maturation, curing, and roasting. *J. Agric. Food Chem.* 51, 4273–4277.
- Comai, L., Young, K., Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Henikoff, S., 2004. Efficient discovery of DNA polymorphisms in natural populations by EcoTilling. *Plant J.* 37, 778–786.
- Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L., Henikoff, S., 2003. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics* 164, 731–740.
- Gruber, P., Becker, W.-M., Hofmann, T., 2005. Influence of the Maillard reaction on the allergenicity of rAra h 2, a recombinant major allergen from peanut. *J. Agric. Food Chem.* 53, 2289–2296.
- Henikoff, S., Comai, L., 2003. Single-nucleotide mutations for plant functional genomics. *Annu. Rev. Plant Biol.* 54, 375–401.
- Henikoff, S., Till, B.J., Comai, L., 2004. TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiol.* 135, 630–636.
- Joseph, L.M., Hymowitz, T., Schmidt, M.A., Herman, E.M., 2006. Evaluation of *Glycine* germplasm for nulls of the immunodominant allergen P34/Gly m Bd 30 k. *Crop Sci.* 46, 1755–1763.
- Kang, I.-H., Gallo, M., 2007. Cloning and characterization of a novel peanut allergen Ara h 3 isoform displaying potentially decreased allergenicity. *Plant Sci.* 172, 345–353.
- Kochert, G., Stalker, H., Gimenes, M., Galgaro, L., Lopes, C., Moore, K., 1996. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *Am. J. Bot.* 83, 1282–1291.
- Lehmann, K., Schweimer, K., Reese, G., Randow, S., Suhr, M., Becker, W.-M., Veiths, S., Rosch, P., 2006. Structure and stability of 2S albumin-type peanut allergens: implications for the severity of peanut allergic reactions. *Biochem. J.* 395, 463–472.
- Maleki, S.J., Chung, S.-Y., Champagne, E.T., Raufman, J.-P., 2000. The effects of roasting on the allergenic properties of peanut proteins. *J. Allergy Clin. Immunol.* 106, 763–768.
- Ozias-Akins, P., 2007. Peanut. In: Pua, E., Davey, M. (Eds.), *Biotechnology in Agriculture and Forestry, Transgenic Crops VI*, vol. 61. Springer-Verlag, Berlin Heidelberg, pp. 81–105.
- Ozias-Akins, P., Ramos, M.L., Chu, Y., 2006. Hypoallergenic foods beyond infant formulas. In: Maleki, S., Burks, A.W., Helm, R.M. (Eds.), *Food Allergy*. ASM Press, Washington, DC, pp. 287–306.
- Ramos, M., Fleming, G., Chu, Y., Akiyama, Y., Gallo, M., Ozias-Akins, P., 2006. Chromosomal and phylogenetic context for conglutin genes in *Arachis* based on genomic sequence. *Mol. Gen. Gen.* 275, 578–592.
- Ramos, M.L., Huntley, J.J., Maleki, S.J., Ozias-Akins, P., 2008. Identification and characterization of a hypoallergenic ortholog of Ara h 2.01. *Plant Mol. Biol.*, doi: 10.1007/s11103-008-9428-z.
- Restani, P., Ballabio, C., Corsini, E., Flocchi, A., Isoardi, P., Magni, C., Poiesi, C., Terracciano, L., Duranti, M., 2005. Identification of the basic subunit of Ara h 3 as the major allergen in a group of children allergic to peanuts. *Ann. Allergy Asthma Immunol.* 94, 262–266.
- Roth, B.M., Pruss, G.J., Vance, V.B., 2004. Plant viral suppressors of RNA silencing. *Virus Res.* 102, 97–108.
- Sen, M., Kopper, R., Pons, L., Abraham, E.D., Burks, A.W., Bannon, G.A., 2002. Protein structure plays a critical role in peanut allergen stability and may determine immunodominant IgE-binding epitopes. *J. Immunol.* 169, 882–887.
- Simpson, C.E., 2001. Use of wild *Arachis* species/introgression of genes into *A. hypogaea* L. *Peanut Sci.* 28, 114–116.
- Stanley, J.S., King, N., Burks, A.W., Huang, S.K., Sampson, H., Cockrell, G., Helm, R.M., West, C.M., Bannon, G.A., 1997. Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2. *Arch. Biochem. Biophys.* 342, 244–253.
- Teuber, S., Beyer, K., Comstock, S., Wallowitz, M., 2006. The big eight foods: clinical and epidemiological overview. In: Maleki, S., Burks, A., Helm, R. (Eds.), *Food Allergy*. ASM Press, Washington, DC, pp. 49–79.
- Vaucheret, H., 2006. Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes Dev.* 20, 759–771.