

Identification and characterization of a hypoallergenic ortholog of Ara h 2.01

M. Laura Ramos · James J. Huntley ·
Soheila J. Maleki · Peggy Ozias-Akins

Received: 27 June 2008 / Accepted: 29 October 2008
© Springer Science+Business Media B.V. 2008

Abstract Peanut (*Arachis hypogaea* L.), can elicit type I allergy becoming the most common cause of fatal food-induced anaphylactic reactions. Strict avoidance is the only effective means of dealing with this allergy. Ara h 2, a peanut seed storage protein, has been identified as the most potent peanut allergen and is recognized by approximately 90% of peanut hypersensitive individuals in the US. Because peanut has limited genetic variation, wild relatives are a good source of genetic diversity. After screening 30 *Arachis duranensis* accessions by EcoTILLing, we characterized five different missense mutations in *ara d 2.01*. None of these polymorphisms induced major conformational modifications. Nevertheless, a polymorphism in the immunodominant epitope #7 (S73T) showed a 56–99% reduction in IgE-binding activity and did not affect T cell epitopes, which must be retained for effective immunotherapy. The identification of natural hypoallergenic

isoforms positively contributes to future immunological and therapeutic studies and peanut cultivar development.

Keywords Peanut allergy · Ara h 2 isoforms · *Arachis duranensis* · EcoTILLING · Hypoallergenic proteins

Introduction

Peanut (*Arachis hypogaea* L.) is a nutritious food whose consumption has been linked to many health benefits. Conversely, peanuts can trigger IgE-mediated food allergy becoming the most common cause of fatal and near-fatal food-induced anaphylactic reactions (Pham and Rudner 2000; Bock et al. 2001; Sampson 2004). Seed storage proteins, which play an important role as sources of nitrogen and amino acids during germination, elicit this kind of allergy. Among these seed storage proteins, Ara h 2 comprises ~6–9% of total seed proteins (Koppelman et al. 2001) and is one of the major reactive peanut allergens recognized by more than 90% of the peanut-allergic individuals in the US (Koppelman et al. 2004). Several approaches including anti-IgE therapy, DNA vaccines and allergen-specific immunotherapy (SIT) or sublingual immunotherapy (SLIT) have been proposed to reduce allergic response, but some treatments are not recommended for use in the case of peanut allergy because of the risk of anaphylaxis (de Leon et al. 2007). To date, there is still no effective treatment other than strict avoidance of peanuts and peanut by-products (Burks et al. 1998). However, widespread use of peanuts in consumer foods makes avoidance difficult. Therefore, it is necessary to explore strategies designed to reduce the ability of food allergens to elicit dangerous allergic responses yet retain

Electronic supplementary material The online version of this article (doi:10.1007/s11103-008-9428-z) contains supplementary material, which is available to authorized users.

M. L. Ramos · P. Ozias-Akins (✉)
Department of Horticulture, University of Georgia Tifton
Campus, Tifton, GA 31793, USA
e-mail: pozias@uga.edu

Present Address:

M. L. Ramos
Department of Plant, Soil and Agricultural Systems, Southern
Illinois University, Carbondale, IL 62901, USA

J. J. Huntley
National Center for Genome Resources, Santa Fe,
NM 87505, USA

S. J. Maleki
USDA-ARS-SRRC, New Orleans, LA 70124, USA

their ability to interact with T cells (de Leon et al. 2007) that respond to redirect the immune system to a tolerant state (Bannon et al. 2001). Scanning mutagenesis of the IgE-reactive allergen peptides revealed that changes in single amino acids could eliminate IgE binding (Stanley et al. 1997; Shin et al. 1998; Rabjohn et al. 2002) reducing the risk of systemic allergic reactions. In this context, the use of a peanut that contains hypoallergenic proteins (low IgE-binding activity and conserved T-cell activating capacity) of potent allergens potentially could blunt the allergic reaction in sensitized individuals who inadvertently consume this food.

As a member of the 2S albumin family, Ara h 2 contains eight cysteines that are conserved in terms of context and position (Shewry et al. 1995; Jose-Estanyol et al. 2004). Ten linear IgE epitopes were identified as distributed throughout the length of the amino acid sequence and three of them (epitopes 3, 6 and 7) were considered to be immunodominant (Stanley et al. 1997). Ara h 2 is synthesized as a precursor with an additional N-terminal 21-amino acid signal peptide that is required for protein transport to storage vacuoles. Based upon high sequence homology to the similar Ara h 6 allergen, Ara h 2 is thought to exist as a four-helix bundle, stabilized by as many as four disulfide linkages (Jose-Estanyol et al. 2004; Lehmann et al. 2006). The disulfide bonds are hypothesized to contribute to the thermal and proteolytic stability of Ara h 2 and may possibly dictate which IgE-binding epitopes are immunodominant (Sen et al. 2002; Barre et al. 2005; Lehmann et al. 2006).

Ara h 2 has two isoforms, Ara h 2.01 (Burks et al. 1992) and Ara h 2.02 which has a duplication of 12 amino acids containing an extra copy of DPYSPS (Chatel et al. 2003). The gene for Ara h 2.01 belongs to the A-subgenome (putative donor *A. duranensis*, $2n = 2x = 20$, AA) and Ara h 2.02 to the B-subgenome (putative donor *A. ipaensis*, $2n = 2x = 20$, BB) of tetraploid peanut ($2n = 4x = 40$, AABB) (Ramos et al. 2006). The orthologs of *ara h 2*, *ara d 2.01* (NCBI accn. EF609641) and *ara i 2.02* (NCBI accn. EF609642), share 100% sequence similarity with *ara h 2.01* and *ara h 2.02*, respectively (Ramos et al. 2006) making it possible to use the biochemical information available from Ara h 2 for evaluation of *A. duranensis* isoforms and vice versa.

In plants, one of the most common methods for suppressing gene expression and eliminating a protein is through RNA interference (RNAi) (Ozias-Akins et al. 2006). This approach has been used to silence the major apple allergen Mal d1 (Gilissen et al. 2005) and is already being applied in peanut (Dodo et al. 2005; Gilissen et al. 2005; Ozias-Akins et al. 2006). Cosuppression also has been used to remove an immunodominant allergen from soybean (Herman et al. 2003). However, these techniques require the production of transgenic plants. As an alternative, TILLING

(Targeting Induced Local Lesions in Genomes) is a technique to identify single nucleotide polymorphisms (SNPs) in an induced mutant population (McCallum et al. 2000). A variation of TILLING, EcoTILLING, has been adapted to the discovery of polymorphisms in natural populations (Henikoff and Comai 2003; Comai et al. 2004). Some SNPs may alter amino acid sequence or truncate proteins thereby affecting gene function and potentially allergenicity.

Since cultivated peanut has relatively little allele sequence variation as opposed to wild relatives which are rich sources of variation, the present work was performed to evaluate natural variation in *ara d 2.01*, the ortholog of *ara h 2.01*. Polymorphisms in the open reading frame as well as in regulatory regions of the gene were identified after screening 30 accessions of *A. duranensis* available from the US germplasm collection. One variant isoform containing a polymorphism (S73T) in the immunodominant epitope #7 showed a reduction of 56–99% in its IgE-binding activity compared with the wild-type Ara d 2.01. The characterization of potentially hypoallergenic protein variants in wild germplasm can guide peanut breeding programs to develop hypoallergenic or reduced allergenicity lines and/or generate materials suitable for immunotherapy.

Materials and methods

Genetic material

Genomic DNA was prepared from 30 *A. duranensis* accessions obtained from the US peanut germplasm collection, Griffin, GA (Supplementary Table 3). Leaf DNA was extracted using the DNeasy 96 plant kit (Qiagen). The DNA concentration was measured with a Packard fluorometer using picogreen (Molecular Probes). Individual and pooled-sample plates were prepared at 0.5 ng/μl final concentration. *Arachis duranensis* accession PI 262133 (AD5), whose *ara d 2.01* sequence (GenBank accn. EF609641) was identical to the A-subgenome sequence in *A. hypogaea*, was designated as wild-type and pooled with each accession sample as control.

Nested PCR primers

PCR 5' primer 813 (5'-GGAGTGAAAAAGAGAAGAGA ATA-3') and 3' primer 817 (5'-TCAAGATGGTTACA ACTCTGCAGCAACA-3') were designed to amplify *ara h 2* genes from both A- and B-subgenomes in peanut. To specifically amplify *ara h 2.01* or its ortholog *ara d 2.01* we used a non-conserved region of the promoter to design the 5' primer, 815 (5'-CGATTTACTCATGTACAATTAA CAATAGAT-3') along with the 3' primer 371 (5'-CAGC AACAAAACATAGACAACGCC-3'). We validated the

ability of primer pair 815/371 to specifically amplify a 1,280 bp fragment from *A. hypogaea* cv Georgia Green and the respective orthologs from the putative diploid progenitors *A. duranensis* and *A. ipaensis*. Only one expected size band of 1,280 bp was amplified from the A-subgenome (Supplementary Fig. 4) and from *A. duranensis* and confirmed by sequence analysis.

High-throughput mutation screening for *ara d 2.01*

The first PCR with primer pair 813/817 was carried out in a 25 µl final volume under the following conditions: 94°C 5 min; 35 cycles at 94°C 30 s, 52°C 30 s, 72°C 2 min; 72°C 5 min; hold at 4°C. An aliquot (2 or 4 µl) from a 1:40 dilution of the first PCR was used as input for a second PCR carried out in 10 µl final volume with infrared fluorescent dye (IRD)-labeled primers (MWG Biotech). The primer cocktail included 24 µl of 100 µM IRD700-labeled primer 815, 16 µl of 100 µM unlabeled primer 815, 32 µl of 100 µM IRD800-labeled primer 371, 8 µl of 100 µM unlabeled primer 371. Touchdown PCR was conducted in a thermal cycler (MJ Research PTC-200 Peltier) as follows: heat denaturation at 95°C for 2 min followed by six cycles of 94°C 30 s, 58°C 30 s decreasing 1°C/cycle, temperature ramp 0.5°C/s to 72°C 1 min 20 s, then 45 cycles of 94°C 30 s, 55°C 30 s with a temperature ramp 0.5°C/s to 72°C, and finally extension, 72°C 7 min. The heteroduplex formation step was performed with a denaturation at 99°C for 10 min, 70 cycles of reannealing at 70°C 20 s each, decreasing 0.3°C/s/cycle with a final hold at 4°C. After PCR amplification, samples (5 µl from the second PCR or 70–100 ng) were digested for 15 min at 45°C in 10 µl total volume with 0.03 µl celery juice prepared as described for Cel I purification (Till et al. 2006). To stop the reaction 5 µl of 0.15 M EDTA were added per sample while keeping the samples on ice. The samples were cleaned using Sephadex G-50–150 medium uniformly loaded in 96-well MultiScreen Plates (Cat. No. MAHVN 45, Millipore) using a Multi-Screen Column loader (Cat. No. MACL 096 45) following the manufacturer's instructions. The samples were collected in a catch plate, transferred to a 96-well PCR plate (#650-PCR, DOT Scientific, Inc.), dried in a centrifugal evaporator (ISS110 Speed Vac System, Thermo Savant) for 15 min at high temperature, resuspended in 8 µl of formamide loading buffer, and heated (Slade et al. 2005). Samples (0.8–1 µl) were loaded on a 48-lane shark-tooth comb and electrophoresed on a 6.5% polyacrylamide gel (Li-Cor) in 1X TBE at 1,500 V, 40 mA, 5°C on a Li-Cor 4300 DNA Analyzer. Images were visually analyzed for presence of cleavage products using Adobe Photoshop (Adobe Systems, Inc.) and Gelbuddy (Zerr and Henikoff 2005).

Ara h 2.01 homology model generation and refinement to represent Ara d 2.01 polymorphisms

An Ara h 2.01 homology model using the Ara h 2.01 amino acid sequence (NCBI Accession AAK96887 (Viquez et al. 2001) was generated using the Swiss-Model homology modeling server (Peitsch 1995; Guex and Peitsch 1997; Schwede et al. 2003). Furthermore, the model was built using the lowest energy Ara h 6 NMR structure as a modeling template (Lehman et al. to be published). Prior to submission of the Ara h 2.01 sequence to the Swiss-Model server, the putative N-terminal signal peptide was removed as there is no corresponding sequence in the Ara h 6 NMR structure, and although likely present in the mature Ara h 2.01 protein, was judged not likely to contribute to the overall Ara h 2.01 fold.

The returned homology model (129 amino acids in length, beginning with arginine 22 and ending with aspartic acid 156 of the corresponding Ara h 2.01 sequence) was further refined in two steps using the AMBER8 suite of molecular dynamic simulation algorithms (Case et al. 2004), a canonical ensemble and the *ff03* force field (Duan et al. 2003). Three disulfide bonds were created between cysteines 45 and 91, 92 and 140, and 106 and 148 using corresponding disulfide bonds in the Ara h 6 structure as a guide. One corresponding disulfide bridge present in the Ara h 6 structure (between cysteines 16 and 73 of our model) was not created as these side chains were over 7 Å distant.

The model was initially energy minimized by performing a steep-descent in vacuo minimization (1,000 steps, conjugate gradient) to relieve bad steric interactions. The model was then charge neutralized and solvated using a TIP3P water box, and a 2 ns solvated molecular dynamic simulation performed by first completing a restrained minimization (keeping the protein fixed) to relieve bad contacts in the surrounding solvent, followed by a unrestrained minimization to relieve bad contacts in the entire system. This was followed by a 20 ps constant-pressure position-restrained dynamic simulation (weak restraints on the protein, raising the temperature from 0 to 300 K) to relax the position of the solvent molecules. The simulation was completed by performing a 2 ns constant-pressure (1 atm.) constant-temperature (300 K) molecular dynamic simulation with isotropic position scaling, Langevin dynamics with a collision frequency of 1.0/ps, and the SHAKE algorithm to constrain hydrogen bond lengths. The resultant refined Ara h 2.01 model structure was then analyzed and validated by PROCHECK (Laskowski et al. 1996).

Models of the *A. duranensis* EcoTILLING variants (Q47E, S73T and D158 N) were constructed using our refined Ara h 2.01 structure and Swiss-Pdb Viewer (Guex

and Peitsch 1997). Model variant structures were analyzed to assess the extent to which amino acid substitution would affect hydrogen bonding, the overall protein fold and the potential for altered IgE binding.

Sequence analysis

Genomic DNA from putative variants was used as input to amplify *ara d 2.01* with 815–371 primers and the high proofreading *Taq* PrimeSTAR (Takara). Sequence analysis was conducted to rank the mutations based on their probability of affecting protein function. The SIFT (Sorting Intolerant from Tolerant) bioinformatic program (Ng and Henikoff 2006) was used to predict the severity of each identified mutation.

SDS-PAGE, Western blots and IgE-binding analysis

Seeds from *A. duranensis* accessions were evaluated for Ara d 2.01 content. Seed protein extraction was carried out as described by Koppelman et al. (2001). Protein quantification was performed by the Bradford method using a Bio-Rad kit and bovine serum albumin (BSA) as the standard. 10 µg of total protein were loaded per lane and separated in 12% SDS-PAGE gels (Sambrook et al. 1989). The proteins were transferred overnight to Immuno-Blot PVDF membranes, 0.2 µm (BioRad Cat. No. 162-0174) using a BioRad electrophoretic transfer system at 30 V. The membrane was blocked with 5% non-fat milk in PBS-T (137 mM NaCl, 2.7 mM KCl, 1.4 mM KH₂PO₄, 4.3 mM Na₂HPO₄ · 12H₂O, 0.1% Tween-20) at room temperature for 1 h with gentle agitation. After blocking, the membrane was probed with primary chicken

anti-Ara h 2 (1:8,000) for 1 h followed by secondary anti-chicken HRP antibody from Sigma (Cat. No. A9046) in 2% non-fat milk in PBS-T (1:100,000) for 1 h. Custom polyclonal antibody was obtained from Sigma after providing them with purified Ara h 2 protein. Semi-quantitative analysis of Ara d 2.01 protein was performed by densitometry with Quantity One (BioRad) and all samples were normalized to the amount of Ara d 2.01/10 µg of total protein in wild-type. The normalized blot was probed with a 1:20 dilution of five different sera from peanut-allergic individuals (NF, SM, HW, DAM and CM) overnight at 4°C. The membrane was washed three times with PBS-T for 5 min after each antibody incubation. ECL Plus detection reagents (Amersham Cat. No. RPN2133) were used for anti-chicken IgG or anti-human IgE (Sigma) conjugated to horseradish peroxidase (HRP) for chemifluorescent or chemiluminescent detection, respectively. The Western blot molecular weight standard MagicMark XP (Invitrogen Cat. No. LC5602) was used because it could be detected by this kit. Further IgE immunoblot analyses were performed in triplicate using 150 ng of total protein.

Results

Identification of natural SNP variants in the ORF of *ara d 2.01*

EcoTILLING of a 1,280 bp amplicon of *ara d 2.01* allowed the identification of SNPs within the ORF in seven (AD1, AD2, AD13, AD18, AD19, AD29 and AD30) out of thirty *A. duranensis* accessions (Supplementary Fig. 1). The

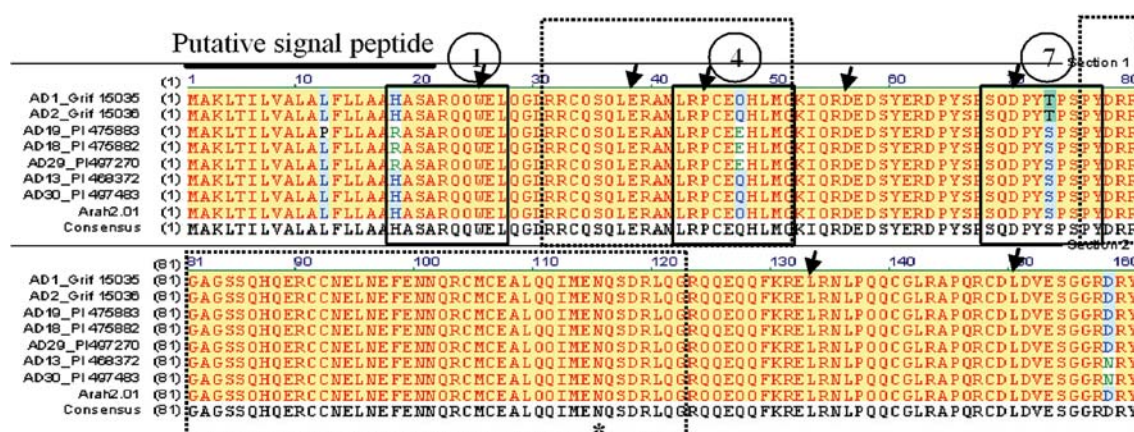


Fig. 1 Predicted amino acid alignment of Ara d 2.01 in *A. duranensis* accessions carrying the missense mutations detected by EcoTILLING. The linear immuno IgE epitopes #1, #4 and #7 affected by amino acid changes H18R, Q47E and S73T are boxed with solid lines. The putative N-terminal 21-amino acid signal peptide contains two amino acid changes L12P and H18R. None of these SNPs

impacted the putative N-glycosylation site (indicated with asterisk). Dashed boxes display the dominant T-cell epitopes reported by Glaspole et al. (2005) and affected by Q47E. None of the missense mutations matched the sites mutated by King et al. (2005) in order to produce their recombinant mutant Ara h 2 (pointed out with arrow heads)

polymorphisms, confirmed by sequence analysis, were determined to correspond to eight SNPs at nucleotide positions 35, 53, 132, 139, 210, 218, 276 and 472 (Supplementary Fig. 2). Alignment of the deduced proteins confirmed that five of the eight SNPs (at positions 35, 53, 132, 218, and 472) resulted in missense mutations (Fig. 1, Table 1). Three missense mutations, H18R (His–Arg), Q47E (Gln–Glu) and S73T (Ser–Thr), are located within experimentally determined IgE epitopes #1, #4 and #7 (Stanley et al. 1997) and T-cell epitopes (Glaspole et al. 2005; King et al. 2005) (Fig. 1, Table 1). The other two missense mutations, L12P (Leu–Pro) and D158N (Asp–Asn), fall in the putative signal peptide and at the end of the protein, respectively. Furthermore, none of these amino acid changes matched with the mutations introduced in the recombinant mutant Ara h 2 (King et al. 2005) (Fig. 1).

An alignment including the deduced amino acid sequence of Ara d 2.01 and the natural variants carrying the missense mutations (L12P, H18R, Q47E, S73T and D158N) was submitted to SIFT (Sorting Intolerant from Tolerant), a bioinformatic program that assigns a severity score to each mutation (Ng and Henikoff 2006). SIFT scores the changes from 0 to 1, where 0 is damaging and 1 is neutral. Scores <0.05 are predicted to be deleterious although there is a 20% false-positive rate. All amino acid changes in the alignment were predicted to be tolerated by Ara d 2.01 but with widely divergent SIFT scores (Table 1). Some of these less tolerated changes might affect the correct protein folding and/or correct targeting or secretion leading to protein degradation. Thus, to test whether any of the accessions lacked Ara d 2.01 as a consequence of carrying a missense mutation, 5 µg of total seed proteins were electrophoresed, blotted, and probed with chicken anti-Ara h 2 IgG serum (Supplementary Fig. 3). Western blot analysis confirmed that there was no absence of Ara d 2.01 expression in these accessions.

Point mutations and indels in non-coding regions

Several indels (insertions and deletions) and point mutations were also effectively detected by EcoTILLING in the 5' and 3' non-coding regions of the *A. duranensis* accessions that also contained SNPs in the ORF. Of the natural variants containing SNPs within their ORFs (Supplementary Table 1, Fig. 1), accessions AD18 and AD19 have a modification in a CAAT box located around –481 bp and AD2 has a SNP in the GARE (GA-Responsive Element) motif at –295 from the start codon. The other accessions have different indels and SNPs but they are not affecting important motifs (Lescot et al. 2002). More accessions displayed putative mutations in non-coding regions but not in their ORFs; therefore, they have not yet been sequenced.

Table 1 Predicted amino acid changes (shown in bold) detected in *ara d 2.01* ORF of *A. duranensis* accessions selected by EcoTILLING and confirmed by nucleotide sequence analysis

SNPs position	aa Change	Nature of the mutation	IgE epitope affected ^a	T-cell epitope affected ^b	SIFT score ^c	<i>A. duranensis</i> accession	Short ID
35	L12P	Apolar → apolar			0.11	PI 475883	AD19
53	H18R	Polar ^d → charged polar	1 AHASARQQWEL ARASARQQWEL		1.00	PI 475882 PI 475883	AD18 AD19
139	Q47E	Polar → charged polar	4 LRPCEQHLMQ LRPCEEHLMQ	RRCQSQLERANLRPCEEHLM RRCQSQLERANLRPCEEHLM	1.00	PI 475882 PI 475883	AD18 AD19
218	S73T	Polar → polar	7 SQDPYSPSY SQDPYTPSY		0.28	PI 497270 Grif 15035	AD29 AD1
472	D158N	Charged polar → polar			0.92	Grif 15036 PI 468372 PI 497483	AD2 AD13 AD30

^a Some isoforms of *ara d 2.01* and *ara h 2.01* share 100% sequence similarity thus the biochemical information from Ara h 2, like IgE epitopes and T-cell epitopes, becomes available for this ortholog

^b T-cell epitope sequence reported by Glaspole et al. (2005)

^c SIFT scores predicted to affect protein function: 0 = damaging, 1 = neutral; only *ara d 2.01* and *ara h 2.01* orthologs/alleles were compared

^d The polarity/charge of histidine is highly dependent on the physiochemical environment surrounding the side chain

Computational modeling of predicted proteins

The refined Ara h 2.01 homology model (Fig. 2a), built using the lowest energy Ara h 6 NMR structure as a modeling template, consists of an alpha helical core stabilized by three disulfide bridges. The amino acid sequence identity between Ara h 6 (NCBI Accession 1W2Q_A, 127 amino acids in length) and the full-length Ara h 2 sequence (NCBI Accession AAK96887, 135 amino acids in length after cleavage of the putative n-terminal signal peptide), is 58% (72/123 identities). Considering the degree of sequence homology between Ara h 2.01 and Ara h 6, the refined Ara h 2.01 homology model is very similar to the Ara h 6 NMR structure (Protein Data Bank ID 1W2Q, Lehman et al. unpublished). The core of the protein is flanked by regions of apparent polypeptide-chain disorder including stretches of the N- and C-termini and an extended loop that encompasses residues Y60 through S84. Analysis of the refined homology by PROCHECK (Laskowski et al. 1996) indicates that 85% of the residues (excluding glycine, proline and end-residues) fall into “most favored regions” and 15% fall into “additionally allowed regions” of the Ramachandran Plot (data not shown).

Our refined Ara h 2.01 homology is remarkably similar to a recently published structure that also used the Ara h 6 NMR structure as a modeling template (Lehmann et al. 2006). We note only minor differences, expectedly in disordered regions of the protein. The Ara h 2.01 homology model reported by Lehmann et al. (2006) does possess an additional core-stabilizing disulfide bridge (four total) between cysteines 16 and 73. We did not elect

to form this particular disulfide linkage as these cysteines were over 7 Å distant in our model. The omission of this particular disulfide bridge, however, does not appear to drastically alter the similarity displayed between the two models.

Effects of mutations on IgE-binding activity

To assess the IgE-binding capacity of the different Ara d 2.01 isoforms, Western blot analysis was performed using seed protein extract from representative accessions carrying SNPs only in *ara d 2.01* ORF. Accessions carrying mutations in the putative signal peptide (variant L12P) were not tested because this mutation does not affect any immunodominant epitope. Total seed protein extracts were normalized for loading equal amounts of Ara d 2.01, as measured by anti-Ara h 2 chicken polyclonal antibody, and tested for binding to serum from five patients with peanut hypersensitivity (NF, SM, HW, DAM and CM) (Supplementary Table 2). Semi-quantitative analysis by densitometry allowed the separation of patient sera into two groups, one with low (SM and DAM) and one with high (NF, HW, and CM) IgE-binding activity (data not shown). For the high IgE-binding activity group, accessions at each extreme, AD12 (containing wild-type protein that displayed the strongest binding of IgE) and AD1 (containing the variant isoform that displayed the weakest reaction) were tested in replicate. The missense mutation S73T (Ser–Thr) carried by AD1 occurs within the immunodominant epitope #7 (SQDPYSPSY/SQDPYTPSY) (Stanley et al. 1997). Triplicate samples of total seed proteins normalized for loading equal amounts of Ara d 2.01 present in AD12 and AD1 were

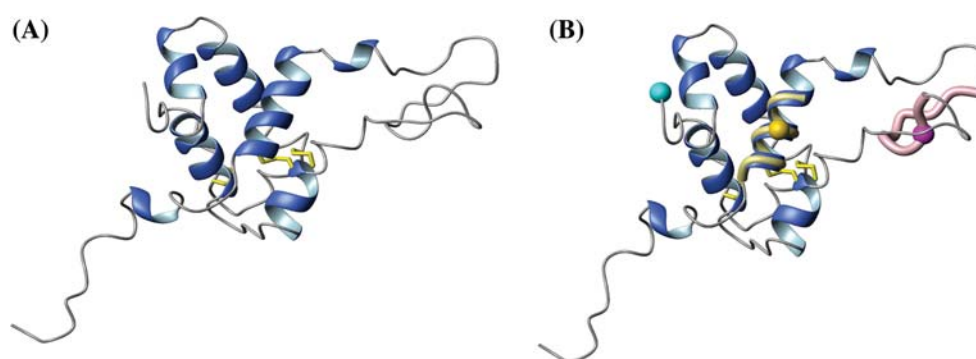


Fig. 2 **a** Refined homology model of Ara h 2.01. Initial coordinates for the model were generated using the Swiss-Model homology modeling server using the NMR structure from Ara h 6 as a modeling template (PDB ID 1W2Q). The structure was further refined by completing a 2 ns solvated and charge-neutralized molecular dynamic simulation using AMBER8. Corresponding disulfide bridges present in the Ara h 6 template structure were created for the Ara h 2 homology model between cysteines 45 and 91, 92 and 140, and 106 and 148 (shown in yellow). **b** Refined homology model of Ara h 2

annotated with the location of observed *A. duranensis* amino acid variation detected by EcoTILLING. The Q47E variant (gold sphere) resides at a position within the corresponding Ara h 2 IgE epitope #4 (indicated by a thicker backbone representation in light gold), while the S73T variant (pink sphere) resides at a position within the corresponding Ara h 2 IgE epitope #7 (thicker backbone representation in light pink). The D158 N variant (cyan sphere) resides at the c-terminus of the model structure and not within any reported Ara h 2 IgE epitope

analyzed by immunoblot with individual sera from three patients. Semi-quantitative analysis by densitometry showed 56% (CM), 80.4% (HW) and 99% (NF) reduction in IgE binding to Ara d 2.01 in AD1 compared with the Ara d 2.01 wild-type protein in AD12 (Fig. 3a–c). Although in CM (Fig. 3a), IgE binding of Ara d 2.01 present in AD1 dropped by ~50% and this drop is visually apparent in the blot, the difference was not statistically significant by densitometry, probably because the background in this membrane was relatively high. Significant reduction was demonstrated with the other two sera. The IgE binding reduction observed with these sera was reconfirmed when 150 ng total seed protein was used (Fig. 4a–c). SM serum, belonging to the low IgE binding group, also displayed reduced IgE-binding at this concentration (Fig. 4d).

Fig. 3 Immunoblot (IgE binding) and densitometric analysis. Individual sera from CM (a), HW (b) and NF (c) belonging to the high IgE-binding group were used to test total seed protein extracts from AD12 and AD1. All patient sera reacted with multiple peanut proteins (obvious in the region of the gel shown in c), but only Ara d 2.01 was analyzed in this experiment. The samples were loaded in triplicate and normalized for Ara d 2.01. Densitometry showed 56.2% (a), 80.4% (b) and 99% (c) reduction in IgE binding capacity for the Ara d 2.01 isoform present in AD1 compared with the control (AD12). The differences observed in b and c, but not a, were statistically significant (letters A and B) for Ara d 2.01. Error bars indicate standard deviation

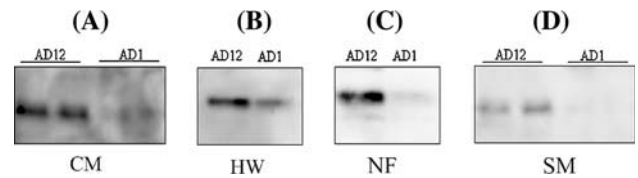
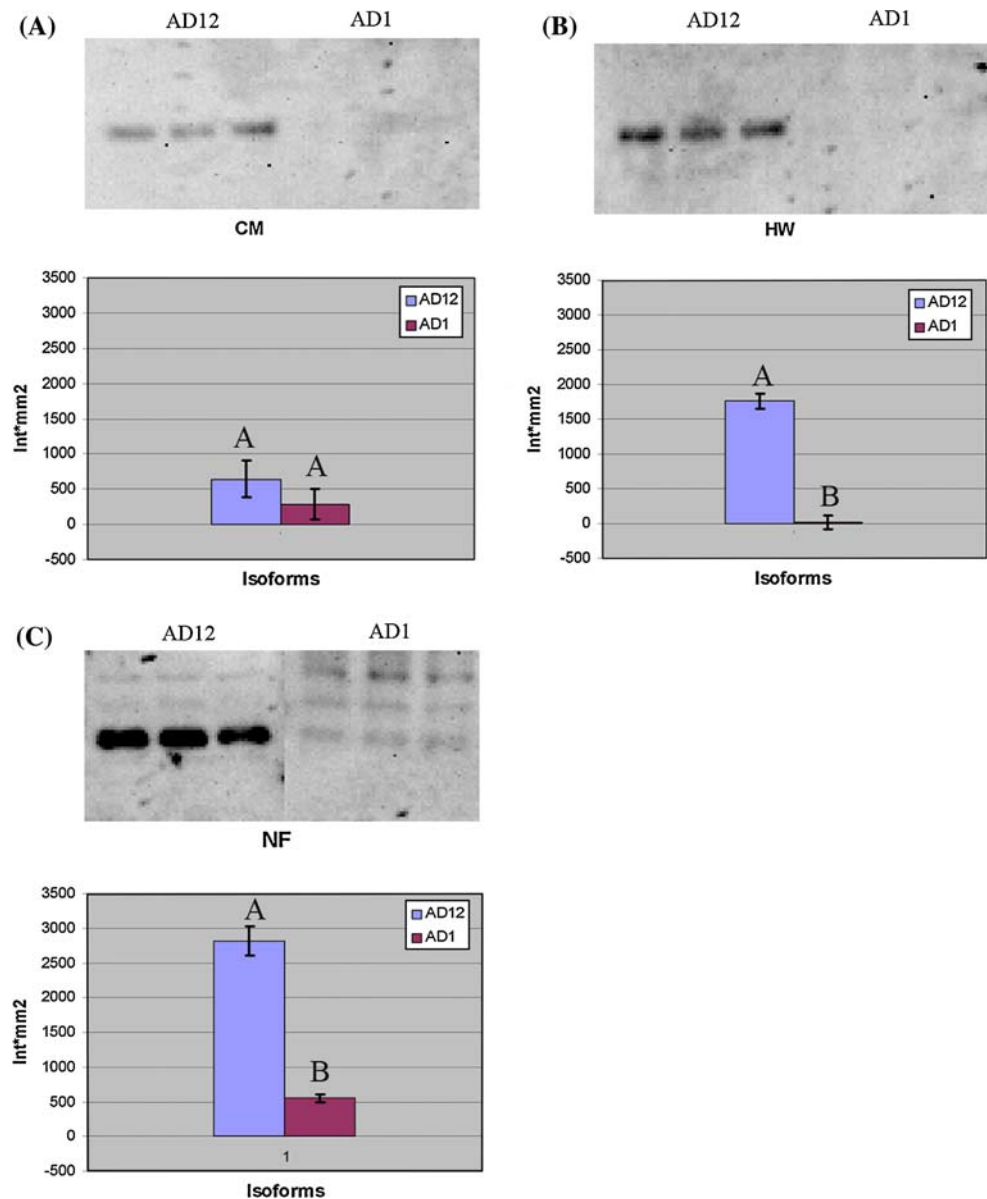


Fig. 4 Immunoblot (IgE binding) to test 150 ng of total seed protein extracts from AD12 and AD1. Individual sera from CM (a), HW (b) and NF (c) belonging to the high IgE-binding group were used. Also, SM (d) a serum from the low IgE-binding group displayed a reduction in IgE binding

Discussion

We have extended the search for hypoallergenic proteins present in wild relatives of peanut with the aim of

identifying and characterizing natural variants of Ara h 2 proteins displaying loss or reduction of IgE-binding capacity. Using EcoTILLING we have analyzed 30 accessions from *A. duranensis*, a wild diploid relative of peanut, and the putative A-subgenome donor. This technique allowed the identification and selection of seven accessions (~23%) displaying mutations in the ORF of *ara d 2.01*. The natural mutation rate found for this particular gene was one SNP/420 bp. The analysis of deduced amino acid sequences showed that five of the eight SNPs resulted in missense mutations and three of them H18R, Q47E and S73T affect IgE and T-cell epitopes. Although H18R has been located in epitope #1 (Stanley et al. 1997), this change together with mutation L12P fall in the putative signal peptide. The cleavage site of the signal peptide in planta has not been experimentally determined for Ara h 2. The most probable site for signal peptide cleavage in Ara h 2 was predicted to be between positions 21 and 22 at the site ASA-RQ (Ramos et al. 2006). Taking into account that Stanley and co-workers (Stanley et al. 1997) were using the derived amino acid sequence of Ara h 2 for the design and synthesis of IgE epitope peptides, we believe that IgE epitope #1 includes part of the signal peptide in its sequence. Thus, the observed IgE binding to this epitope could be a consequence of the recognition of sequences that correspond to the immature protein or to nonspecific cross-reaction of IgE against these amino acids. Furthermore, the amino acid changes L12R and H18R are not affecting the hydrophobic properties of the signal peptide suggesting the correct subcellular targeting of the proteins carrying these modifications, avoiding degradation that could be detected as lack of expression. The mutation positioned near the end of the protein, D158N, does not impact any of the linear IgE epitopes previously described (Stanley et al. 1997) nor the regions that appear to contain the majority of the peptides that interact with T-cells from most patients (Glaspole et al. 2005; King et al. 2005).

Furthermore, none of the SNPs knocked out *ara d 2.01* expression, affected the putative *N*-glycosylation site (position 115, Fig. 1) or the cysteine residues. Cysteines are known to be of paramount importance in disulfide bridge formation and therefore are responsible for the correct three-dimensional conformation of the protein. The addition of a cysteine residue in P34/Gly m Bd 30 k, an immunodominant human allergen in soybean, for example, induces a misfolded protein making it inactive (Herman et al. 2003; Joseph et al. 2006). Also, hypoallergenic variants have been generated by site-directed mutagenesis of selected cysteine residues (Smith and Chapman 1996; Drew et al. 2004). In our refined Ara h 2.01 homology model the omission of one disulfide bridge between cysteines 16 and 73 does not appear to affect the three-dimensional structure of the protein. Despite the

importance of the cysteine residues and the low probability to find a natural variant that carries a mutation at this particular amino acid, it could be interesting to find Ara h 2 mutants in EMS-mutagenized populations with one or more altered disulfide bridges to empirically determine the effect on protein stability and IgE binding activity.

Annotation of the refined Ara h 2.01 homology model with the location of observed amino acid variation between Ara h 2.01 and Ara d 2.01 in *A. duranensis* EcoTILLING variants is shown in Fig. 2b. We have also annotated the model with the location of Ara h 2 IgE epitopes impacted by these variants (Stanley et al. 1997). The Q47-glutamic acid substitution occurs within Ara h 2 IgE epitope #4, residing central to one of the core alpha helices. The substitution imparts a negative charge at this position relative to Ara h 2.01, and may possibly result in a new hydrogen bond formed between the carboxylic acid side chain of E47 and the NH₂ side chain of R43. This particular variant under denaturing conditions did not display significantly reduced IgE binding from sera of peanut-allergic individuals. Because of its location within the helix, and the nature of the substitution, it is not likely that this mutation would result in significant structural changes. Nonetheless, this position was shown to be necessary for efficient IgE binding based upon alanine-scanning mutagenesis studies (Stanley et al. 1997). As this particular study was carried out on synthesized mutant peptides, and not the complete Ara h 2 protein, this may be an indication for the propensity of IgEs from peanut-allergic individuals to recognize this epitope in a linear peptide and not within the context of flanking amino acids.

Substitution of S73 with a threonine occurs within Ara h 2 IgE epitope #7. Based upon our model, epitope #7 resides on the extended, disordered loop segment between residues Y60 through S84. As noted above, this variant did display marked reduction in IgE binding from peanut-allergic individuals. This is somewhat surprising as the substitution of a serine with a threonine is not expected to result in significant alterations to the physiochemical properties at this specific position within the epitope. While this substitution does not significantly change the physiochemical landscape of the surrounding region, our analyses suggest that such a mutation may result in a loss of one hydrogen bond between the carboxylic acid side chain of D70 and the backbone amide of Y72. However, this would not account for the change in reactivity of the linear epitope. Interestingly, earlier immunological analysis of synthesized Ara h 2 peptides by alanine-scanning mutagenesis noted that while not critical for efficient IgE binding, mutation of this serine to an alanine, resulted in an *increase* in IgE binding (Stanley et al. 1997). A more recent study by King et al. (2005) did examine alterations to IgE binding for alanine mutants of complete recombinant form(s) of the allergen,

but did not mutate this particular position. A detailed structural analysis of our model suggests that a S73A mutation is also accompanied by a loss of a hydrogen bond, this time between the side chain at position 73 and the carboxylic acid side chain of D70. These observations imply that there may be subtle structural factors important in the differential IgE recognition of this particular epitope, but testing of this hypothesis would require analysis of properly folded proteins. It is also possible, as previously mentioned, that flanking sequences may play a role in IgE epitope recognition.

Although reduction in IgE-binding capacity does not necessarily translate into reduced allergenic potency, removal of B-cell epitopes is needed to abrogate IgE-mediated adverse reactions and dominant T-cell epitopes must be retained to target T cells for effective immunotherapy (Rolland et al. 2000). Recombinant proteins have been engineered to generate hypoallergenic peanut proteins for Ara h 1, Ara h 2 and Ara h 3 (Burks et al. 1999; Bannon et al. 2001; Rabjohn et al. 2002; King et al. 2005) with properties favorable for immunotherapy. The fact that the natural Ara d 2.01 isoform carrying the modification S73T displays reduction between 56 and 99% in the IgE binding capacity without modification of dominant T cell epitopes responsible for Ara h 2 T-cell antigenicity (Glaspole et al. 2005; King et al. 2005) could be considered as indirect evidence that this Ara d 2.01 isoform has not altered its T-cell capacity. Therefore, after performing more analyses concerning T-cell proliferation, this natural variant could potentially be used for immunotherapy.

However, variation in the degree of reduction in IgE binding has been reported for the engineered recombinant peanut allergens. In almost all cases, some of the peanut-allergic individuals still retained high IgE binding activity, similar to what we observed. Therefore, it has been suggested that the multiple sensitivities displayed by peanut-allergic individuals is a consequence of the promiscuity of IgE binding, making the use of selected allergenic peanut proteins for immunotherapy more difficult (Lewis et al. 2005).

None of the *A. duranensis* accessions with ORF-based mutations showed a lack of Ara d 2.01 expression suggesting that these natural variants, including the one with reduced IgE binding, might be able to retain their native function, an important feature from the perspective of the plant. Ara h 2 is an abundant and integral seed storage protein thus it may be necessary for the hypoallergenic isoforms to retain as much of the native function, properties and three-dimensional structure as possible to avoid deleterious effects on plant growth and development, although silencing experiments suggest that Ara h 2 may not be essential (Chu et al. 2008). The natural Ara d 2.01 variant (S73T) that shows reduced IgE-binding capacity

had not been identified using in vitro mutagenesis; therefore, this new variant could be valuable for therapeutic studies or for further studying the impact of allergen sequence and structure on allergenic properties. In addition, the *A. duranensis* accessions carrying this variant allele might be useful for peanut cultivar development by conventional breeding bypassing the controversial process of genetic modification. Furthermore, the tracking of the trait will be facilitated by the presence of the SNPs which can be targeted specifically by marker assisted selection. In conclusion, natural variation for Ara h 2 isoforms in a wild relative of peanut may be neutral for the plant, but could have utility for clinical applications or breeding a less allergenic peanut.

Acknowledgments This work was supported by the Consortium for Plant Biotechnology Research, The Georgia Peanut Commission, and the Peanut Foundation. We thank Evelyn Morgan, Anne Bell, Hsiapo Cheng, Ye Chu, and Paola Faustinelli for technical assistance.

References

- Bannon GA, Cockrell G, Connaughton C, West CM, Helm R, Stanley JS, King N, Rabjohn P, Sampson HA, Burks AW (2001) Engineering, characterization and in vitro efficacy of the major peanut allergens for use in immunotherapy. *Int Arch Allergy Immunol* 124:70–72. doi:[10.1159/000053672](https://doi.org/10.1159/000053672)
- Barre A, Borges J-P, Culerrier R, Rouge P (2005) Homology modelling of the major peanut allergen Ara h 2 and surface mapping of IgE-binding epitopes. *Immunol Lett* 100:153–158. doi:[10.1016/j.imlet.2005.03.014](https://doi.org/10.1016/j.imlet.2005.03.014)
- Bock SA, Munoz-Furlong A, Sampson HA (2001) Fatalities due to anaphylactic reactions to foods. *J Allergy Clin Immunol* 107:191–193. doi:[10.1067/mai.2001.112031](https://doi.org/10.1067/mai.2001.112031)
- Burks AW, Williams LW, Connaughton C, Cockrell G, O'Brien TJ, Helm RM (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. doi:[10.1016/0091-6749\(92\)90469-I](https://doi.org/10.1016/0091-6749(92)90469-I)
- Burks W, Sampson H, Bannon G (1998) Peanut allergens. *Allergy* 53:725–730
- Burks AW, King N, Bannon GA (1999) Modification of a major peanut allergen leads to loss of IgE binding. *Int Arch Allergy Immunol* 118:313–314. doi:[10.1159/000024114](https://doi.org/10.1159/000024114)
- Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA, Crowley M, Brozell S, Tsui V, Gohlke H, Mongan J, Hornak V, Cui G, Beroza P, Shafmeister C, Caldwell JW, Ross WS, Kollman PA (2004) AMBER 8. University of California, San Francisco
- Chatel J-M, Bernard H, Orson FM (2003) Isolation and characterization of two complete Ara h 2 isoforms cDNA. *Int Arch Allergy Immunol* 131:14–18. doi:[10.1159/000070429](https://doi.org/10.1159/000070429)
- Chu Y, Faustinelli P, Ramos ML, Hajdich M, Stevenson S, Thelen JJ, Maleki SJ, 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*. <http://dx.doi.org/10.1021/jf802600r>
- Comai L, Young K, Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, Henikoff S (2004)

- Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J* 37:778–786. doi:[10.1111/j.0960-7412.2003.01999.x](https://doi.org/10.1111/j.0960-7412.2003.01999.x)
- de Leon MP, Rolland JM, O’Hehir RE (2007) The peanut allergy epidemic: allergen molecular characterisation and prospects for specific therapy. *Expert Rev Mol Med* 9:1–18. doi:[10.1017/S1462399407000208](https://doi.org/10.1017/S1462399407000208)
- Dodo H, Konan K, Viquez O (2005) A genetic engineering strategy to eliminate peanut allergy. *Curr Allergy Asthma Rep* 5:67–73. doi:[10.1007/s11882-005-0058-0](https://doi.org/10.1007/s11882-005-0058-0)
- Drew AC, Eusebius NP, Kenins L, de Silva HD, Suphioglu C, Rolland JM, O’Hehir RE (2004) Hypoallergenic variants of the major latex allergen Hev b 6.01 retaining human T lymphocyte reactivity. *J Immunol* 173:5872–5879
- Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* 24:1999–2012. doi:[10.1002/jcc.10349](https://doi.org/10.1002/jcc.10349)
- Gilissen LJWJ, Bolhaar STHP, Matos CI, Rouwendal GJA, Boone MJ, Krens FA, Zuidmeer L, van Leeuwen A, Akkerdaas J, Hoffmann-Sommergruber K, Knulst AC, Bosch D, van de Weg WE, van Ree R (2005) Silencing the major apple allergen Mal d 1 by using the RNA interference approach. *J Allergy Clin Immunol* 115:364–369. doi:[10.1016/j.jaci.2004.10.014](https://doi.org/10.1016/j.jaci.2004.10.014)
- Glaspole IN, de Leon MP, Rolland JM, O’Hehir RE (2005) Characterization of the T-cell epitopes of a major peanut allergen, Ara h 2. *Allergy* 60:35–40. doi:[10.1111/j.1398-9995.2004.00608.x](https://doi.org/10.1111/j.1398-9995.2004.00608.x)
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714–2723. doi:[10.1002/elps.1150181505](https://doi.org/10.1002/elps.1150181505)
- Henikoff S, Comai L (2003) Single-nucleotide mutations for plant functional genomics. *Annu Rev Plant Biol* 54:375–401. doi:[10.1146/annurev.arplant.54.031902.135009](https://doi.org/10.1146/annurev.arplant.54.031902.135009)
- Herman EM, Helm RM, Jung R, Kinney AJ (2003) Genetic modification removes an immunodominant allergen from soybean. *Plant Physiol* 132:36–43. doi:[10.1104/pp.103.021865](https://doi.org/10.1104/pp.103.021865)
- Jose-Estanyol M, Gomis-Ruth FX, Puigdomenech P (2004) The eight-cysteine motif, a versatile structure in plant proteins. *Plant Physiol Biochem* 42:355–365. doi:[10.1016/j.plaphy.2004.03.009](https://doi.org/10.1016/j.plaphy.2004.03.009)
- Joseph LM, Hymowitz T, Schmidt MA, Herman EM (2006) Evaluation of glycine germplasm for nulls of the immunodominant allergen P34/Gly m Bd 30 k. *Crop Sci* 46:1755–1763. doi:[10.2135/cropsci2005.12-0500](https://doi.org/10.2135/cropsci2005.12-0500)
- King N, Ricki HJ, Steven S, Stefan V, Dirk L, Lina H, Hugh S, Laurent P, Wesley B, Gary AB (2005) Allergenic characteristics of a modified peanut allergen. *Mol Nutr Food Res* 49:963–971. doi:[10.1002/mnfr.200500073](https://doi.org/10.1002/mnfr.200500073)
- Koppelman SJ, Vlooswijk RAA, Knippels LMJ, Hessing M, Knol EF, van Reijssen FC, Bruijnzeel-Koomen CAFM (2001) Quantification of major peanut allergens Ara h 1 and Ara h 2 in the peanut varieties Runner, Spanish, Virginia, and Valencia, bred in different parts of the world. *Allergy* 56:132–137. doi:[10.1034/j.1398-9995.2001.056002132.x](https://doi.org/10.1034/j.1398-9995.2001.056002132.x)
- Koppelman SJ, Wensing M, Ertmann M, Knulst AC, Knol EF (2004) Relevance of Ara h1, Ara h2 and Ara h3 in peanut-allergic patients, as determined by immunoglobulin E Western blotting, basophil histamine release and intracutaneous testing: Ara h2 is the most important peanut allergen. *Clin Exp Allergy* 34:583–590. doi:[10.1111/j.1365-2222.2004.1923.x](https://doi.org/10.1111/j.1365-2222.2004.1923.x)
- Laskowski R, Rullmann J, MacArthur M, Kaptein R, Thornton J (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8:477–486. doi:[10.1007/BF00228148](https://doi.org/10.1007/BF00228148)
- Lehmann K, Schweimer K, Reese G, Randow S, Suhr M, Becker W-M, Vieths S, Rasch P (2006) Structure and stability of 2S albumin-type peanut allergens: implications for the severity of peanut allergic reactions. *Biochem J* 395:463–472. doi:[10.1042/BJ20051728](https://doi.org/10.1042/BJ20051728)
- Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 30:325–327. doi:[10.1093/nar/30.1.325](https://doi.org/10.1093/nar/30.1.325)
- Lewis SA, Grimshaw KEC, Warner JO, Hourihane JOB (2005) The promiscuity of immunoglobulin E binding to peanut allergens, as determined by Western blotting, correlates with the severity of clinical symptoms. *Clin Exp Allergy* 35:767–773. doi:[10.1111/j.1365-2222.2005.02522.x](https://doi.org/10.1111/j.1365-2222.2005.02522.x)
- McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiol* 123:439–442. doi:[10.1104/pp.123.2.439](https://doi.org/10.1104/pp.123.2.439)
- Ng PC, Henikoff S (2006) Predicting the effects of amino acid substitutions on protein function. *Annu Rev Genomics Hum Genet* 7:61–80. doi:[10.1146/annurev.genom.7.080505.115630](https://doi.org/10.1146/annurev.genom.7.080505.115630)
- Ozias-Akins P, Ramos ML, Chu Y (2006) Hypoallergenic foods beyond infant formulas. In: Maleki SJ, Burks AW, Helm RM (eds) Food allergy: comprehension, treatment and prevention. ASM Press, Herndon, VA
- Peitsch MC (1995) Protein modeling by e-mail. *Nat Biotechnol* 13:658–660. doi:[10.1038/nbt0795-658](https://doi.org/10.1038/nbt0795-658)
- Pham T, Rudner E (2000) Peanut allergy. *Cutis* 65:285–289
- Rabjohn P, West CM, Connaughton C, Sampson HA, Helm RM, Burks AW, Bannon GA (2002) Modification of peanut allergen Ara h 3: effects on IgE binding and T cell stimulation. *Int Arch Allergy Immunol* 128:15–23. doi:[10.1159/000057999](https://doi.org/10.1159/000057999)
- 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 Genet* 275:578–592
- Rolland JM, Douglass J, O’Hehir RE (2000) Allergen immunotherapy: current and new therapeutic strategies. *Expert Opin Investig Drugs* 9:515–527. doi:[10.1517/13543784.9.3.515](https://doi.org/10.1517/13543784.9.3.515)
- Sambrook J, Fritsch E, Maniatis T (1989) Molecular cloning, a laboratory manual, 2nd edn. Cold Spring Harbor Press, Cold Spring Harbor, NY
- Sampson HA (2004) Update on food allergy. *J Allergy Clin Immunol* 113:805–819. doi:[10.1016/j.jaci.2004.03.014](https://doi.org/10.1016/j.jaci.2004.03.014)
- Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* 31:3381–3385. doi:[10.1093/nar/gkg520](https://doi.org/10.1093/nar/gkg520)
- Sen M, Kopper R, Pons L, Abraham EC, Burks AW, Bannon GA (2002) Protein structure plays a critical role in peanut allergen stability and may determine immunodominant IgE-binding epitopes. *J Immunol* 169:882–887
- Shewry PR, Napier JA, Tatham AS (1995) Seed storage proteins: structures and biosynthesis. *Plant Cell* 7:945–956
- Shin DS, Compadre CM, Maleki SJ, Kopper RA, Sampson H, Huang SK, Burks AW, Bannon GA (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. doi:[10.1074/jbc.273.22.13753](https://doi.org/10.1074/jbc.273.22.13753)
- Slade AJ, Fuerstenberg SI, Loeffler D, Steine MN, Facciotti D (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat Biotechnol* 23:75–81. doi:[10.1038/nbt1043](https://doi.org/10.1038/nbt1043)
- Smith AM, Chapman MD (1996) Reduction in IgE binding to allergen variants generated by site-directed mutagenesis: contribution of disulfide bonds to the antigenic structure of the major

- house dust mite allergen Der p 2. *Mol Immunol* 33:399–405. doi: [10.1016/0161-5890\(95\)00150-6](https://doi.org/10.1016/0161-5890(95)00150-6)
- Stanley JS, King N, Burks AW, Huang SK, Sampson H, Cockrell G, Helm RM, West CM, Bannon GA (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. doi: [10.1006/abbi.1997.9998](https://doi.org/10.1006/abbi.1997.9998)
- Till BJ, Zerr T, Comai L, Henikoff S (2006) A protocol for TILLING and Ecotilling in plants and animals. *Nat Protoc* 1:2465–2477. doi: [10.1038/nprot.2006.329](https://doi.org/10.1038/nprot.2006.329)
- Viquez OM, Summer CG, Dodo HW (2001) Isolation and molecular characterization of the first genomic clone of a major peanut allergen, Ara h 2. *J Allergy Clin Immunol* 107:713–717. doi: [10.1067/mai.2001.113522](https://doi.org/10.1067/mai.2001.113522)
- Zerr T, Henikoff S (2005) Automated band mapping in electrophoretic gel images using background information. *Nucleic Acids Res* 33:2806–2812. doi: [10.1093/nar/gki580](https://doi.org/10.1093/nar/gki580)

Supplementary Table 1. SNPs and indels in 5' and 3' non-coding regions of *A. duranensis* accessions confirmed by sequence analysis. *In silico* analysis of the promoter motifs was performed using PlantCARE

5' Region							CAAT BOX ^a (127-132)										GARE ^b motif							
nucleotide	29	49	54	67	96	97	127	129	158	170	180	218	219	230	312	316	353/372	379	394	484	598			
AD1																								
AD2							CAA ^c		AT ^c								G/T	C/G	# ^{cd}					
AD 13		G/A	G/A	T/C			C/T		ACCTA ^c				A/G		GA ^c		C/G		C ^c A/C					
AD 18		T/C			C/A	A/G	A/G		A/T		T ^c		A/G		GA ^c			G/T						
AD 19		G/T		T/C		A/G		A/G		A/T		T ^c		A/G		GA ^c			G ^c		G/T			
AD29																								
AD 30		T/C			A/G				ACCTA #				A/G		GA ^c			G/T						

3'UTR							
nucleotide	1099	1110	1123	1158	1164	1186	1190
AD1	A ^c		T ^c				
AD2						TA/CC	G/A
AD 13		A ^c			T ^c		
AD 18					T ^c		
AD 19					T ^c		
AD29							
AD 30		A ^c		ATAT ^c	T ^c		

^a CAAT box: common *cis*-acting element in promoter and enhancer regions; ^bGARE motif: gibberellin-responsive element

^cIndel: insertion/ deletion; ^dACGACATAACAACATGGCA.

Supplementary Table 2. Patient/serum information. All sera were obtained from patients with either a convincing history of allergy to peanut or a positive peanut challenge. All sera were obtained with the consent of each patient and the approval of the institutional review board of University of Arkansas for Medical Sciences.

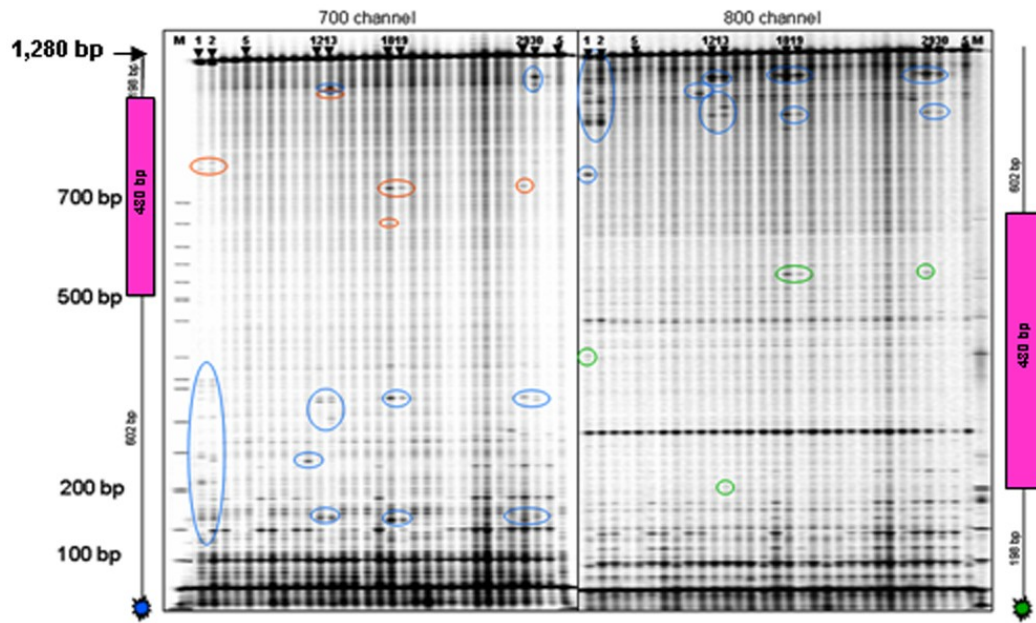
Patient	Birth date	Rast	Screen/ Class	Gender	Other allergy
DAM	1972	>100	4+	F	soy, tree nuts, beans, etc.
NF	09/18/91	>100	4+	M	unknown
HW	?	>100	4+	M	unknown
SM	01/02/65	>100	2+	M	tree nuts
CM	02/15/78	70	3+	F	tree nuts

Supplementary Table 3. *Arachis duranensis* accessions obtained from the US peanut germplasm collection, Griffin, GA and analyzed by Ecotilling for *ara d 2.01* variants.

Accession	Plant ID	Country, state	Short ID*
Grif 15035	WiSVg1510-A	Paraguay, Boqueron	AD1
Grif 15036	WiSVg1510-B	Paraguay, Boqueron	AD2
Grif 15037	WiSVgJsQ 1506-E	Paraguay, Boqueron	AD3
Grif 15038	WiSVgJsQ 1506-W	Paraguay, Boqueron	AD4
PI 262133	10038	Argentina	AD5
PI 468197	GKBSPSc 30060	Argentina	AD6
PI 468200	GKBSPSc 30064	Argentina	AD7
PI 468201	GKBSPSc 30065	Argentina	AD8
PI 468203	GKBSPSc 30064	Argentina	AD10
PI 468319	GKBSPSc 30073	Bolivia	AD11
PI 468320	GKBSPSc 30074	Bolivia	AD12
PI 468372	ScBo 15101	Paraguay	AD13
PI 475844	GKBSPSc 30069	Bolivia	AD14
PI 475845	GKBSPSc 30070	Bolivia	AD15
PI 475846	GKBSPSc 30071	Bolivia	AD16
PI 475847	GKBSPSc 30072	Bolivia	AD17
PI 475882	KSBScC 36002	Argentina	AD18
PI 475883	KSBScC 36003	Argentina	AD19
PI 475884	KSBScC 36004	Argentina	AD20
PI 475885	KSBScC 36005	Argentina	AD21
PI 475886	KSBScC 36006	Argentina	AD22
PI 475887	KSSc 36036	Argentina	AD23
PI 497263	ScVa 21764	Argentina	AD24
PI 497265	ScVa 21767	Argentina	AD25
PI 497266	KSSc 38900	Argentina	AD26
PI 497267	KSSc 38903	Argentina	AD27
PI 497268	KSSc 38904	Argentina	AD28
PI 497270	KSSc 38906	Argentina	AD29
PI 497483	KSSc 38901	Bolivia	AD30
PI 497484	KSSc 38902	Bolivia	AD31

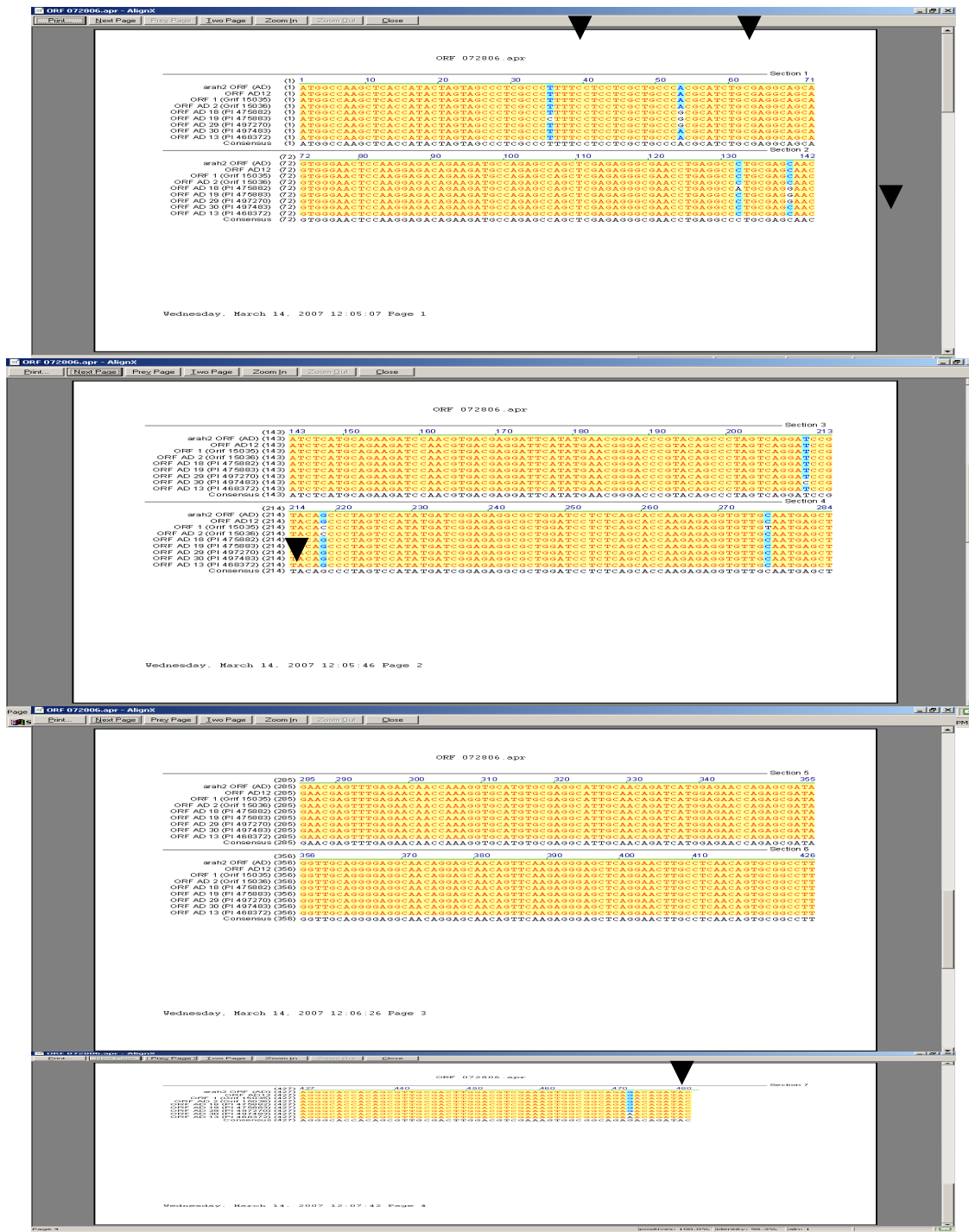
*A short ID was assigned in our lab to facilitate labeling and identification of the sample

Supplementary Fig. 1



The image on the left corresponds to the IRD700 channel, and on the right to the IRD800. Cel I cleaved products appear as darker bands. These bands are shorter than the full-length PCR product and indicate the presence of a putative mutation. The bands circled in red or in green point out putative mutations in the ORF. The bands circled in blue correspond to SNPs or indels detected in the 5' or 3' non-coding regions. Arrow heads designate the wild-type accession (AD5) and accessions with SNPs in the ORF (AD1, AD2 AD13, AD18, AD19, AD 29 and AD30). Upstream sequence: 602 bp, ORF: 480 bp, 3'UTR: 198 bp.

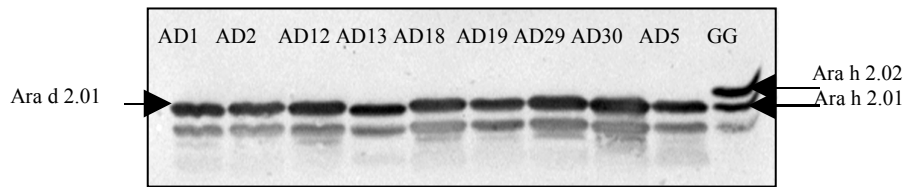
Supplementary Fig. 2



SNP detection within the *ara d 2.01* ORF confirmed by sequence analysis in *A. duranensis* mutants detected by Ecotilling.

SNPs at nucleotide positions 35, 53, 139, 218 and 472 inducing amino changes are pointed out by arrow heads. SNPs at positions 132, 210 and 276 do not change amino acid composition. AD12 was incorporated in this alignment to show that this accession shares 100% sequence identity with *ara d 2.01* from AD5 (PI 262133), the accession taken as wild-type DNA during Ecotilling analysis.

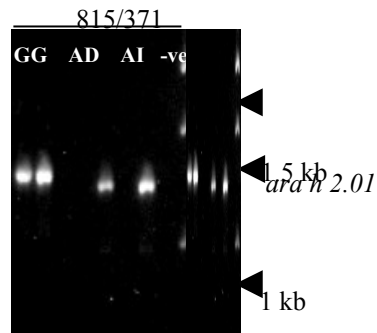
Supplementary Fig. 3



Western blot with chicken anti-Ara h 2 IgG antibody. All accessions showed expression of Ara d 2.01 isoforms. GG: *A. hypogaea* Georgia Green (tetraploid)

AD12 (PI 468320) and AD5 (PI 262133) contain Ara d 2.01 wild-type isoforms. The fainter band below Ara h 2.01 is due to cross-reactivity of the Ara h 2 antibody with Ara h 6.

Supplementary Fig. 4



PCR product analysis for homeolog-specific TILLING primers performed on 2% Nu-Sieve GTG agarose. Primers 815/371 were used to specifically amplify one band corresponding to *ara h 2.01* from *A. hypogaea* cv Georgia Green (GG) or its ortholog from *A. duranensis* (AD), no band was amplified from *A. ipaensis* (AI) as was expected. -ve: negative.