Published in final edited form as:
J Allergy Clin Immunol. 2012 May ; 129(5): 1321–1328.e5. doi:10.1016/j.jaci.2012.02.012.

A bioinformatics approach to identify patients with symptomatic peanut allergy using peptide microarray immunoassay

Jing Lin, PhD[a,*], **Francesca M. Bruni, MD**[b,*], **Zhiyan Fu, PhD**[a], **Jennifer Maloney, MD**[a], **Ludmilla Bardina, MSc**[a], **Attilio L. Boner, MD**[b], **Gustavo Gimenez, BSc**[a], and **Hugh A. Sampson, MD**[a]

[a]Pediatric Allergy and Immunology, Mount Sinai School of Medicine, New York

[b]the Pediatric Department, University of Verona, Verona

Abstract

Background—Peanut allergy is relatively common, typically permanent, and often severe. Double-blind, placebo-controlled food challenge is considered the gold standard for the diagnosis of food allergy–related disorders. However, the complexity and potential of double-blind, placebo-controlled food challenge to cause life-threatening allergic reactions affects its clinical application. A laboratory test that could accurately diagnose symptomatic peanut allergy would greatly facilitate clinical practice.

Objective—We sought to develop an allergy diagnostic method that could correctly predict symptomatic peanut allergy by using peptide microarray immunoassays and bioinformatic methods.

Methods—Microarray immunoassays were performed by using the sera from 62 patients (31 with symptomatic peanut allergy and 31 who had outgrown their peanut allergy or were sensitized but were clinically tolerant to peanut). Specific IgE and IgG_4 binding to 419 overlapping peptides (15 mers, 3 offset) covering the amino acid sequences of Ara h 1, Ara h 2, and Ara h 3 were measured by using a peptide microarray immunoassay. Bioinformatic methods were applied for data analysis.

Results—Individuals with peanut allergy showed significantly greater IgE binding and broader epitope diversity than did peanut-tolerant individuals. No significant difference in IgG_4 binding was found between groups. By using machine learning methods, 4 peptide biomarkers were identified and prediction models that can predict the outcome of double-blind, placebo-controlled food challenges with high accuracy were developed by using a combination of the biomarkers.

Conclusions—In this study, we developed a novel diagnostic approach that can predict peanut allergy with high accuracy by combining the results of a peptide microarray immunoassay and bioinformatic methods. Further studies are needed to validate the efficacy of this assay in clinical practice.

Corresponding author: Jing Lin, PhD, Pediatric Allergy and Immunology, Mount Sinai School of Medicine, Box 1198, New York, NY 10029. jing.s.lin@mssm.edu.
*These authors contributed equally to this work.

Disclosure of potential conflict of interest: H.A. Sampson is a consultant for Allertein Therapeutics, LLC; is on the advisory board for ImmusanT and Novartis; has received research support from the Food Allergy Initiative (FAI) and the National Institutes of Health, National Institute of Allergy and Infectious Diseases; is a consultant and scientific advisor for the FAI; is a medical advisor for the Food Allergy and Anaphylaxis Network; is a scientific advisor for the University of Nebraska FARRP; and is 45% owner of Herbal Springs, LLC. The rest of the authors declare that they have no relevant conflicts of interest.

Keywords

Epitope mapping; peptide microarray; peanut allergy; bioinformatics; machine learning; allergy diagnosis; epitope biomarker

Peanut (*Arachis hypogea*) allergy is one of the most common food allergies, and recent studies from North America and the United Kingdom place the prevalence of peanut allergy among children at slightly more than 1%.[1-3] As opposed to many other childhood food allergies (milk, egg, soy, etc) that most patients outgrow, peanut allergy tends to persist throughout life with only approximately 20% of the children allergic to peanuts "outgrowing" their allergy.[4-6]

Many peanut allergens (designated Ara h 1-11)[7] have been characterized, and sequential IgE-binding epitopes have been defined for some.[8-12] Several studies have shown in both peanut-allergic patients and sensitized animals that the majority of the IgE response is directed at Ara h 1, Ara h 2, Ara h 3, and Ara h 6.[13,14] Previous studies mapping IgE epitopes on Ara h 1, Ara h 2, and Ara h 3 have found a positive correlation between IgE epitope diversity, as represented by the number of epitopes recognized, and the severity of patients' clinical reactivity.[12,15]

Double-blind, placebo-controlled food challenges (DBPCFCs) are the "gold standard" for diagnosing food allergy,[16] but they are time-consuming, expensive, and stressful for the patient. In addition, they place the patient at risk for a potentially life-threatening anaphylactic reaction.[17] The skin prick test (SPT) has also been used to diagnose clinical peanut allergy.[18,19] For example, Sporik et al[19] have reported a SPT wheal diameter of 8 mm or more to be 100% specific in predicting positive challenges to peanut in young children attending an Allergy Clinic in Melbourne, Australia. However, other studies have not found this value to be very predictive.[20] Several studies have evaluated the measurement of serum food specific-IgE (sIgE) levels as a method for predicting the outcome of a food challenge.[21,22] In particular, a peanut sIgE level of 15 kU_A/L was found to have a more than 95% predictive value for a positive oral peanut challenge in a study of children with atopic dermatitis.[16] However, large numbers of patients still require oral food challenges, because their peanut sIgE level or SPT wheal diameter is below the established diagnostic decision levels. In addition, high levels of sIgE or a larger wheal diameter do not necessarily associate with clinical allergy while patients with very low sIgE levels or a small wheal diameter may have severe allergic reactions.

As a novel epitope mapping tool, the peptide microarray immunoassay has been utilized more frequently in the past few years and applied in several studies of milk and peanut allergy.[12,15,23-25] Compared with SPTand the measurement of allergen sIgE levels, determining epitope sIgE levels by using peptide microarrays appears to provide more information about the patient's clinical condition and thus may be more useful in the diagnosis/prognosis of food allergy. However, similar to SPT and the measurement of allergen sIgE levels, the epitope-binding patterns are heterogeneous between patients and there is an overlap between epitope sIgE values from different subject groups.[9,12] A combination of several epitope biomarkers may be more informative. Bioinformatic methods, such as supervised machine learning and classification, may be useful in identifying epitope biomarkers and developing prediction models by using a combination of these biomarkers.

The purpose of our study was to use the peptide microarray immunoassay to compare IgE and IgG_4 binding to the peptides of 3 major peanut allergens (Ara h 1, Ara h 2, and Ara h 3)

between patients with symptomatic peanut allergy and patients who had outgrown their allergy or were sensitized but clinically tolerant to peanut ingestion. More importantly, we sought to develop a novel allergy diagnostic method that could correctly identify patients with symptomatic peanut allergy by applying bioinformatic analyses on peptide microarray data.

Methods

Study population

Sixty-two children and adolescents were recruited selectively from a larger group of patients referred to the Mount Sinai Allergy Clinic for the evaluation of peanut allergy by using DBPCFC between 2001 and 2007. All study subjects had an evaluation that consisted of an extensive history, physical examination, SPT, and measurement of serum peanut sIgE levels. Patients with a positive history of reacting to peanut, patients who had avoided peanut in the diet because of a positive family history of peanut allergy, patients who had a positive SPT to peanut, and/or patients with a sIgE level of 0.35 kU_A/L or more to peanut were recruited. Patients with relatively low serum peanut sIgE levels obtained within 1 year of performing the DBPCFC were selected. Eleven nonatopic individuals (no history of allergic symptoms, negative serum peanut sIgE level) were included in the study as negative controls. All parents gave written informed consent before enrollment in the study. The study was reviewed and approved by Mount Sinai's Institutional Review Board.

Peanut sIgE levels, SPTs, DBPCFCs, and Immuno Solid-phase Allergen Chip (ISAC)

Peanut sIgE levels were determined for each sample by using Phadia UniCAP system according to manufacturer's instructions (Phadia, Uppsala, Sweden). sIgE level of more than 0.35 kU_A/L was considered positive. SPTs and DBPCFCs were performed as previously described.[26-29] Serum sIgE levels to peanut component allergens—Ara h 1, Ara h 2, and Ara h 3— were determined by using ISAC (ImmunoCAP ISAC Assay kit IgE, Phadia) according to manufacturer's instructions. More details can be found in this article's Online Repository at www.jacionline.org.

Peptide microarray

A library of overlapping peptides, consisting of 15 amino acids with an offset of 3, corresponding to the amino acid sequence of Ara h 1, Ara h 2, and Ara h 3, was printed in 2 sets of duplicates onto Arrayit SuperEpoxy glass slides (Arrayit Corporation, Sunnyvale, Calif) as previously described.[15]

Immunolabeling was performed as previously described with some modifications.[30,31] In brief, the slides were blocked with 400 μL of 1% human serum albumin (HSA) in PBS containing 0.05% Tween 20 (PBS-T) for 60 minutes at room temperature, then incubated with 250 μL of patient serum diluted 1:5 in PBS-T/HSA for 24 hours at 4°C, followed by incubation for 24 hours at 4°C with a cocktail of several mAbs including 1 monoclonal anti-human IgG_4-fluorescein isothiocyanate (SouthernBiotech, Birmingham, Ala), diluted 1:500, and3 monoclonal biotinylated anti-human IgE antibodies: 1 from Invitrogen (Carlsbad, Calif), diluted 1:250; 1 from BD Biosciences Pharmingen (San Jose, Calif), diluted 1:250; and 1 as a gift from Phadia, biotinylated in our laboratory and diluted 1:1000 in PBS-T/HSA. Slides were then incubated for 3 hours at 31°C with a cocktail of Anti-Biotin_Dendrimer_Oyster 550 (350) and Anti-fluorescein isothiocyanate_ Dendrimer_Oyster 650 (350) in Dendrimer buffer (Genisphere, Hatfield, Pa), both at 0.6 μg/mL with the addition of 0.02 μg/mL of salmon sperm DNA (Invitrogen), followed by a wash with PBS-T, 15 mM Tris, 0.1× PBS, and 0.05× PBS. Slides were centrifuge dried and

then scanned by using a ScanArrayGx (PerkinElmer, Waltham, Mass). Images were saved as TIF files.

Data analysis

The fluorescence signal of each spot was digitized with the Scan Array Express Microarray Analysis System (PerkinElmer), exported as comma-delimited files, and transformed into robust Z scores, as previously described with minor modifications.[31] In brief, Z scores of each peptide were calibrated (for IgE only) by subtracting the median Z score of the same peptide of the negative controls run in the same experiment and smoothed by using the weighted average of the peptide itself and the 2 flanking peptides, by using the formula $Z = 0.25Z_{-1} + 0.5Z_0 + 0.25Z_{+1}$. An individual peptide is considered positive if its weighted average Z score is more than 3, meaning that the signal was above the background with *P* value of less than .003.

Comparison of IgE and IgG_4 binding diversity was done by using the Mann-Whitney test. IgE and IgG_4 binding epitopes were identified by using TileMap, a tool for tiling array analysis,[32] on the basis of a hierarchical empirical Bayes model with moving average method.[33] As hundreds of peptides were analyzed simultaneously, the false discovery rate was calculated to adjust for multiple comparisons.

Supervised machine learning was employed to determine the boundary between allergic and tolerant groups as follows: IgE reactions (represented as Z scores) to all peptides were imported into the LNKnet software (MIT Lincoln Laboratory, Lexington, Mass). Two machine learning methods—decision tree[34] and support vector machine[35,36]—were used separately to develop/train the prediction models and select a combination of the least number of peptides with the highest accuracy in classifying patients. The parameters of the machine learning methods were trained and tuned through an iterative process, and 5-fold cross-validation was used to evaluate the prediction performances. The end products are the prediction models for patient classification that are based on IgE reactions to a selected set of peptide biomarkers. The prediction models could be applied for new patient classification.

The performance of the method was represented as overall accuracy Q2 = p/N, where p is the total number of correctly predicted cases and N is the total number of cases. In addition, the diagnostic performance of the machine learning method was evaluated and compared with that of other methods by using receiver operating characteristics (ROC) curves.[37,38] Details of the machine learning methods are posted in this article's Online Repository at www.jacionline.org.

Results

Demographic and clinical characteristics of the 62 patients included in this study are listed in Table I. During the oral challenge, 31 patients reacted to peanut (defined as peanut-allergic group) and 31 did not react (defined as peanut-tolerant group). Thirteen of the 31 peanut-tolerant patients had a convincing history of a previous acute reaction following peanut consumption, but they apparently outgrew their allergy, whereas the remaining 18 patients had positive peanut sIgE levels (16 of them had a positive SPT result) but never had clinical reactions following peanut consumption. Peanut-allergic individuals had significantly larger wheal diameters ($P < .0001$) than did peanut-tolerant subjects, but their peanut sIgE levels were not significantly different. Overall, all the 62 patients reacted to at least 1 peptide of Ara h 1, Ara h 2, and Ara h 3 with either IgE or IgG_4 antibodies or both antibodies. IgE and IgG_4 binding to all individual peptides from peanut-tolerant and peanut-

allergic subjects was visualized by using heatmap shown in Fig E1 (see this article's Online Repository at www.jacionline.org).

IgE binding and identification of informative IgE epitopes

Both peanut-allergic and peanut-tolerant individuals had positive IgE binding to peptides of Ara h 1, Ara h 2, and Ara h 3. However, peanut-allergic individuals bound significantly more peptides (median IgE peptides bound = 78 vs 15; $P < .01$) than did peanut-tolerant ones, especially for Ara h 2 (median IgE peptides bound = 10 vs 0; $P < .001$) (Fig 1). Most peanut-allergic individuals (84%) showed positive binding to peptides of Ara h 2, whereas more than half of peanut-tolerant ones (65%) showed no binding.

In addition to broader epitope diversity, peanut-allergic individuals showed significantly more intense IgE binding to peanut epitopes than did peanut-tolerant ones. Thirty-one IgE-binding regions containing multiple epitopes (Ara h 1, 14 regions; Ara h 2, 4 regions; Ara h 3, 13 regions) were identified by using TileMap (Fig 2). These epitopes were informative epitopes; they had significantly greater IgE binding ($P < .01$; false discovery rate < 0.01) in the clinically allergic individuals compared with both tolerant and nonatopic groups. A comparison of the informative epitopes with the epitopes and immunodominant regions identified in previous studies[8,11,12,39] is indicated in Fig 2.

Some of the above-identified binding regions/epitopes were selected for a peptide inhibition assay to test the specificity of IgE binding. As shown in Fig E2 in this article's Online Repository at www.jacionline.org, the complete or partial inhibition of IgE binding by the same peptide indicated that the detected binding was due to epitope sIgE antibodies.

IgG_4 binding

Compared with the IgE binding from peanut-allergic individuals, all individuals had relatively weaker IgG_4 binding to peanut peptides (Fig 2); 24% and 29% of the IgE positive peptides from peanut-tolerant and peanut-allergic groups, respectively, were bound by IgG_4 antibodies. No significant differences in either the diversity or the intensity of IgG_4 binding between peanut-allergic and peanut-tolerant groups were observed (Figs 1 and 2). However, peanut-allergic patients tended to bind more pep-tides with IgG_4 (median IgG_4 peptides bound = 27 vs 13; $P = .17$). Similar to IgE, the difference in IgG_4 binding diversity between groups was more obvious with Ara h 2 peptides (median IgG_4 peptides bound = 4 vs 0; $P = .026$) (Fig 1).

A comparison of the patients who outgrew peanut allergy (n = 13) and the sensitized patients who never had clinical reactions to peanut (n = 18) showed no significant differences between their IgG_4-binding patterns (median IgG_4 peptides bound = 13 vs 16; $P = .76$) (Fig E1).

Peptide biomarkers identified by using the machine learning method

Within the identified informative epitopes, epitopes corresponding to Ara h 2 peptide numbers 8 to 10 had the highest classification power, but it reached only approximately 70% sensitivity and approximately 80% specificity in distinguishing between peanut-allergic and peanut-tolerant groups. A machine learning method was used to identify a combination of peptides that complemented each other to increase the overall diagnostic accuracy. IgG_4 data were not incorporated into the prediction models since no significant differences in IgG_4 binding were observed between peanut-allergic and peanut-tolerant groups.

Three peptides—Ara h 2 peptide number 10 (Ara h 2_10, AA28-42), Ara h 2 peptide number 18 (Ara h 2_18, AA52-66), and Ara h 1 peptide number 16 (Ara h 1_16, AA46-60)

—were identified as peptide biomarkers by using both machine learning methods, decision tree and support vector machine. For support vector machine, 1 additional peptide biomarker —Ara h 3 peptide number 140 (Ara h 3_140, AA418-432)—was added. Among the identified peptide biomarkers, Ara h 2_10 and Ara h 2_18, each demonstrating more than 70% sensitivity and 60% specificity, were the 2 key biomarkers. The other 2 biomarkers—Ara h 1_16 and Ara h 3_140—bound by only a few patients, were selected by the models because they increased the prediction performance when combined with the key Ara h 2 biomarkers. When the 4 biomarkers were combined in the prediction models, the overall accuracy in 5-fold cross-validation reached more than 90% with approximately 90% sensitivity and approximately 95% specificity (Table II).

The developed decision tree, one of the end products of machine learning analysis, is shown in Fig 3: 62 individuals were sorted on the basis of their IgE reactions to the selected peptide biomarkers. Thirty of 31(96.8%) peanut-allergic and 29 of 31 (93.5%) peanut-tolerant individuals could be correctly identified by using this decision tree.

Comparison of the bioinformatic allergy diagnostic method with other methods

The diagnostic performance of this allergy test method was compared with that of other methods (SPT and UniCAP) by using ROC curves (Fig 4),[38,40] which display the relationship between sensitivity and 1–specificity at different thresholds. The area under the ROC curve indicates the diagnostic accuracy/performance of the test methods; the greater the area under the curve, the more precise the method. The diagnostic performances of the peptide microarray data without machine learning analysis, such as IgE-binding diversity and epitope sIgE levels, are also presented in order to demonstrate the usefulness of machine learning.

As shown in Fig 4, the prediction model developed by using machine learning methods has the best diagnostic performance. Although Ara h 2_10 was the key peptide biomarker with the highest predictive power, the measurement of IgE binding to Ara h 2_10 had lower diagnostic performance than did the developed model combining 4 peptide biomarkers. The measurement of IgE binding to Ara h 2 had similar diagnostic performance as SPT and was better than the measurement of IgE binding to Ara h 1 or Ara h 3. Peanut sIgE level measurements by UniCAP had very low predictive power in this cohort, which is close to a random guess (diagonal line). This is because only patients with relatively low peanut sIgE levels were included and there is no significant difference in the peanut sIgE levels between the 2 groups.

In addition, comparison of this approach with the ImmunoCAP ISAC data of 54 patients (87% of study subjects) showed that the prediction model had higher diagnostic sensitivity (90% vs 74%) with similar specificity (see this article's Online Repository at www.jacionline.org).

Discussion

Currently, the most definitive diagnostic test for peanut allergy is the DBPCFC.[16] However, the complexity and potential of DBPCFC to cause life-threatening allergic reactions affects its clinical application.[17] Several studies have investigated the possibility of measuring the the sIgE level to peanut allergens in predicting the outcome of an oral food challenge.[21,22] For example, a study in United Kingdom using a component-resolved diagnostic assay found that children with peanut allergy had a higher response to Ara h 1, Ara h 2, and Ara h 3, especially Ara h 2, than did tolerant children, suggesting that the measurement of IgE binding to Ara h 2 may be a useful tool in predicting peanut allergy.[41] Similar findings were reported in a recent study in which the measurement of Ara h 2 sIgE levels instead of peanut

sIgE levels could correctly distinguish between most peanut-allergic and peanut-tolerant patients.[42]

Compared with the component-resolved diagnostics using peanut allergens, the measurement of IgE specific for allergenic epitopes provides more insight into the molecular basis of the antibody-allergen interaction, which reflects the patients' allergic status and may become a more accurate allergy diagnostic tool. Previous studies have found that IgE diversity to sequential epitopes of Ara h 1, Ara h 2, and Ara h 3 correlated with clinical severity and IgE binding to certain epitopes was significantly different between children with and without symptomatic peanut allergy.[9,15]

In this study, we compared IgE and IgG_4 binding to sequential epitopes of Ara h 1, Ara h 2, and Ara h 3 in 31 patients with symptomatic peanut allergy and 31 clinically tolerant patients who had outgrown their allergy or who were sensitized but never had clinical reactions. We sought to identify peptide biomarkers for the development of a more accurate allergy diagnostic method. IgE and IgG_4 binding to peptides from Ara h 1, Ara h 2, and Ara h 3 was measured simultaneously by using 2 distinct fluorescence channels. As there were largely excessive amounts of each peptide printed onto the slides, the competition effect between IgE and IgG_4 was not significant.

Sensitivity of the peptide microarray platform

In previous studies of peanut allergy,[8,11,12,39] IgE-binding epitopes in patients with a low peanut sIgE level, especially patients with an sIgE level of less than 0.35 kU_A/L, have not been studied because of the limitation of previous detection systems. However, patients with peanut sIgE levels of less than 0.35 kU_A/L may still be clinically reactive. Therefore, we focused on patients with suspected peanut allergy and relatively low peanut sIgE levels in this study.

To increase the sensitivity for patients with very low sIgE levels, we modified the original immunolabeling method and added 1 more step of incubation with a Dendrimer amplifier. The improved platform was capable of detecting IgE binding from the highly diluted serum pool with sIgE levels as low as 0.068 kU_A/L, while binding to negative control remained negative (see Fig E3 in this article's Online Repository at www.jacionline.org). In this study, IgE binding to peanut epitopes was detectable by using the upgraded peptide microarray in 9 of the 10 patients with sIgE levels of less than 0.35 kU_A/L.

IgG_4 binding

Studies on peanut-allergic patients who underwent immunotherapy have found that immunotherapy induces a progressive epitope-specific polyclonal shift from IgE to IgG_4.[43,44] Decreased epitope binding by IgE and a concurrent increase in corresponding epitope binding by IgG_4 has also been reported in patients attaining tolerance to milk allergy.[45] We therefore expected greater IgG_4 binding in the tolerant group. However, both peanut-allergic and peanut-tolerant patients appeared to produce low levels of IgG_4 antibodies to IgE-bound peptides, and there was no significant difference in IgG_4 binding between peanut-allergic and peanut-tolerant groups. As reported in a recently published grass pollen immunotherapy study,[46] allergen specific IgG_4 levels increased significantly during immunotherapy but decreased to near preimmunotherapy baseline levels 2 years after discontinuing immunotherapy, although the subjects remained clinically tolerant. This is consistent with the low IgG_4 binding in the peanut-tolerant group.

Comparison of the epitopes identified between studies

Although most of the previously identified epitopes and immunodominant epitopes overlapped with the informative epitopes identified in this study, there are some discrepancies between them. One reason is the increased sensitivity of the current assay. By using the dendrimer detection system, we identified several new epitopes not found in previous studies, especially in Ara h 3, which had only 4 epitopes reported previously.

The second reason is the difference in the patient group selected as negative controls. Previously reported epitopes were identified by comparing either peanut-sensitive or peanut-allergic patients with nonatopic individuals. IgE binding from peanut-sensitive patients who can tolerate peanut was not considered. In this study, we included peanut-sensitized but peanut-tolerant patients as negative controls and demonstrated (Fig 2) that they also had binding to some peanut peptides. As a result, several previously identified epitopes (ie, epitopes at the c-terminal end of Ara h 1) were not regarded as informative epitopes in this article, because they were bound by both peanut-allergic and peanut-tolerant patients with similar levels of IgE and had no clinical relevance to symptomatic peanut allergy.

Bioinformatic analysis

We used TileMap, a bioinformatics software,[32] to identify informative epitopes by comparing the peanut-allergic individuals with both peanut-tolerant and negative individuals. TileMap was developed for DNA tiling microarray to identify binding sites of transcription factors across the genome by using overlapping probes. It is more advanced and sensitive than other commonly used statistical tests because it takes into account the overlapping feature of the spotted probes and significantly increased the sensitivity of the analysis.

The machine learning method, another bioinformatics tool, was applied to develop prediction models and identify the combination of peptide biomarkers that maximized diagnostic/prediction power. Fivefold cross-validation was used to avoid overfitting the prediction model. Since the patients' data used for performance assessment were unseen during the training phase of machine learning (see this article's Online Repository at www.jacionline.org), the overall accuracy and ROC curves (Table II and Fig 4) can be considered an unbiased assessment of the prediction models and they might be generalized to other independent patient groups. The prediction models are expected to achieve high accuracy in allergy diagnosis for new patients not included in this study.

Importance of Ara h 2 and its peptides

The importance of Ara h 2 over other peanut allergens in predicting clinical sensitivity to peanut has been demonstrated in previous studies with various test platforms.[9,14,41,42] Our findings were in good agreement with those studies and showed that for both IgE and IgG_4, the differences in epitope diversity between peanut-allergic and peanut-tolerant groups were more obvious in Ara h 2. SIgE bindings to Ara h 1 and Ara h 3 were less predictive of clinical allergy, as demonstrated by using ROC curves.

Epitopes of Ara h 2 also have higher predictive power than do those of Ara h 1 and Ara h 3. Two key peptide biomarkers with the highest predictive power were identified in Ara h 2. These 2 peptides mainly overlapped with the 2 previously identified immunodominant IgE-binding epitopes by Stanley et al,[8] which was recognized by all the allergic patients tested. In a subsequent study conducted by Beyer et al,[9] these 2 immunodominant epitopes in Ara h 2 were recognized by most of the patients with symptomatic peanut allergy but by less than 10% of the peanut-tolerant patients and demonstrated the most striking difference within the 8 immunodominant epitopes in Ara h 1, Ara h 2, and Ara h 3.

In addition, preliminary data from our collaborative project on patients treated with peanut oral immunotherapy showed that 2 regions corresponding to these 2 peptide biomarkers were immunodominant in this cohort of peanut-allergic patients. Twenty-two of 23 peanut-allergic patients recognized at least 1 of the 2 regions and 19 recognized both regions. For immunotherapy-treated subjects (n = 12), the binding patterns at these 2 regions shifted from IgE to IgG_4 over time (manuscript in preparation). This confirms the usefulness of the identified peptide biomarkers in the diagnosis of clinical peanut allergy and prediction of tolerance.

In conclusion, we developed an allergy diagnostic approach with high accuracy by combining highly sensitive peptide microarray immunoassays and bioinformatics. Since the patients included in this study had relatively low peanut sIgE levels, further investigation in a larger study population with a wider range of peanut sIgE levels is required before applying this approach for general clinical allergy testing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

H.A.S. was funded in part by grants from the National Institutes of Health/National Institute of Allergy and Infectious Diseases (grant nos. AI44236 and AI066738) and the National Center for Research Resources (grant no. RR026134).

References

1. Sicherer SH, Munoz-Furlong A, Godbold JH, Sampson HA. US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. J Allergy Clin Immunol. 2010; 125:1322–6. [PubMed: 20462634]
2. Burks AW. Peanut allergy. Lancet. 2008; 371:1538–46. [PubMed: 18456104]
3. Sicherer SH, Sampson HA. Peanut allergy: emerging concepts and approaches for an apparent epidemic. J Allergy Clin Immunol. 2007; 120:491–503. [PubMed: 17689596]
4. Hourihane JO, Roberts SA, Warner JO. Resolution of peanut allergy: case-control study. BMJ. 1998; 316:1271–5. [PubMed: 9554896]
5. Hourihane JO. Peanut allergy: recent advances and unresolved issues. J R Soc Med. 1997; 90:40–4. [PubMed: 9176127]
6. Skolnick HS, Conover-Walker MK, Koerner CB, Sampson HA, Burks W, Wood RA. The natural history of peanut allergy. J Allergy Clin Immunol. 2001; 107:367–74. [PubMed: 11174206]
7. Allergen nomenclature approved by the World Health Organization and International Union of Immunological Societies (WHO/IUIS) Allergen Nomenclature Sub-committee. [Accessed 5, 2012] Available at: http://www.allergen.org
8. Stanley JS, King N, Burks AW, Huang SK, Sampson H, Cockrell G, et al. Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2. Arch Biochem Biophys. 1997; 342:244–53. [PubMed: 9186485]
9. Beyer K, Ellman-Grunther L, Jarvinen KM, Wood RA, Hourihane J, Sampson HA. Measurement of peptide-specific IgE as an additional tool in identifying patients with clinical reactivity to peanuts. J Allergy Clin Immunol. 2003; 112:202–7. [PubMed: 12847500]
10. Burks AW, Shin D, Cockrell G, Stanley JS, Helm RM, Bannon GA. Mapping and mutational analysis of the IgE-binding epitopes on Ara h 1, a legume vicilin protein and a major allergen in peanut hypersensitivity. Eur J Biochem. 1997; 245:334–9. [PubMed: 9151961]
11. Rabjohn P, Helm EM, Stanley JS, West CM, Sampson HA, Burks AW, et al. Molecular cloning and epitope analysis of the peanut allergen Ara h 3. J Clin Invest. 1999; 103:535–42. [PubMed: 10021462]

12. Shreffler WG, Beyer K, Chu THT, Burks AW, Sampson HA. Microarray immunoassay: association of clinical history, in vitro IgE function, and heterogeneity of allergenic peanut epitopes. J Allergy Clin Immunol. 2004; 113:776–82. [PubMed: 15100687]
13. van Wijk F, Hartgring S, Koppelman SJ, Pieters R, Knippels LMJ. Mixed antibody and T cell responses to peanut and the peanut allergens Ara h 1, Ara h 2, Ara h 3 and Ara h 6 in an oral sensitization model. Clin Exp Allergy. 2004; 34:1422–8. [PubMed: 15347376]
14. Koppelman SJ, Wensing M, Ertmann M, Knulst AC, Knol EF. 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. 2004; 34:583–90. [PubMed: 15080811]
15. Flinterman AE, Knol EF, Lencer DA, Bardina L, Jager C, Lin J, et al. Peanut epitopes for IgE and IgG_4 in peanut-sensitized children in relation to severity of peanut allergy. J Allergy Clin Immunol. 2008; 121:737–43. [PubMed: 18234310]
16. Sampson HA, Ho DG. Relationship between food-specific IgE concentrations and the risk of positive food challenges in children and adolescents. J Allergy Clin Immunol. 1997; 100:444–51. [PubMed: 9338535]
17. Perry TT, Matsui EC, Conover-Walker MK, Wood RA. Risk of oral food challenges. J Allergy Clin Immunol. 2004; 114:1164–8. [PubMed: 15536426]
18. Clark AT, Ewan PW. Interpretation of tests for nut allergy in one thousand patients, in relation to allergy or tolerance. Clin Exp Allergy. 2003; 33:1041–5. [PubMed: 12911776]
19. Sporik R, Hill DJ, Hosking CS. Specificity of allergen skin testing in predicting positive open food challenges to milk, egg and peanut in children. Clin Exp Allergy. 2000; 30:1540–6. [PubMed: 11069561]
20. Pucar F, Kagan R, Lim H, Clarke AE. Peanut challenge: a retrospective study of 140 patients. Clin Exp Allergy. 2001; 31:40–6. [PubMed: 11167949]
21. Perry TT, Matsui EC, Conover-Walker MK, Wood RA. The relationship of allergen-specific IgE levels and oral food challenge outcome. J Allergy Clin Immunol. 2004; 114:144–9. [PubMed: 15241358]
22. Sicherer SH, Burks AW, Sampson HA. Clinical features of acute allergic reactions to peanut and tree nuts in children. Pediatrics. 1998; 102:e6. [PubMed: 9651458]
23. Cerecedo I, Zamora J, Shreffler WG, Lin J, Bardina L, Dieguez MC, et al. Mapping of the IgE and IgG_4 sequential epitopes of milk allergens with a peptide microarray-based immunoassay. J Allergy Clin Immunol. 2008; 122:589–94. [PubMed: 18774394]
24. Wang J, Lin J, Bardina L, Goldis M, Nowak-Wegrzyn A, Shreffler WG, et al. Correlation of IgE/IgG_4 milk epitopes and affinity of milk-specific IgE antibodies with different phenotypes of clinical milk allergy. J Allergy Clin Immunol. 2010; 125:695–702. [PubMed: 20226304]
25. Lin J, Sampson HA. The role of immunoglobulin E-binding epitopes in the characterization of food allergy. Curr Opin Allergy Clin Immunol. 2009; 9:357–63. [PubMed: 19568005]
26. Sampson HA. Comparative study of commercial food antigen extracts for the diagnosis of food hypersensitivity. J Allergy Clin Immunol. 1988; 82:718–26. [PubMed: 2461401]
27. Bock SA, Sampson HA, Atkins FM, Zeiger RS, Lehrer S, Sachs M, et al. Double-blind, placebo-controlled food challenge (DBPCFC) as an official procedure—a manual. J Allergy Clin Immunol. 1988; 82:986–97. [PubMed: 3060514]
28. Sampson HA, Albergo R. Comparison of results of skin-tests, RAST, and double-blind, placebo-controlled food challenges in children with atopic dermatitis. J Allergy Clin Immunol. 1984; 74:26–33. [PubMed: 6547461]
29. Sicherer SH, Morrow EH, Sampson HA. Dose-response in double-blind, placebo-controlled oral food challenges in children with atopic dermatitis. J Allergy Clin Immunol. 2000; 105:582–6. [PubMed: 10719311]
30. Lin J, Bardina L, Shreffler WG, Andreae DA, Ge YC, Wang JL, et al. Development of a novel peptide microarray for large-scale epitope mapping of food allergens. J Allergy Clin Immunol. 2009; 124:315–22. [PubMed: 19577281]

31. Ayuso R, Sanchez-Garcia S, Lin J, Fu Z, Ibanez MD, Carrillo T, et al. Greater epitope recognition of shrimp allergens by children than by adults suggests that shrimp sensitization decreases with age. J Allergy Clin Immunol. 2010; 125:1286–93. [PubMed: 20471069]
32. Ji HK, Wong WH. TileMap: create chromosomal map of tiling array hybridizations. Bioinformatics. 2005; 21:3629–36. [PubMed: 16046496]
33. Arima S, Lin J, Pecora V, Tardella V. A Bayesian hierarchical model for identifying epitopes in peptide microarray data. Biostatistics. 2012; 13:101–12. [PubMed: 21856651]
34. Pirooznia M, Yang JY, Yang MQ, Deng YP. A comparative study of different machine learning methods on microarray gene expression data. BMC Genomics. 2008; 9:S13. [PubMed: 18366602]
35. Abeel T, Helleputte T, Van de Peer Y, Dupont P, Saeys Y. Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. Bioinformatics. 2010; 26:392–8. [PubMed: 19942583]
36. Noble WS. What is a support vector machine? Nat Biotechnol. 2006; 24:1565–7. [PubMed: 17160063]
37. Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic-tests and procedures—principles and applications. Ann Inter Med. 1981; 94:557–92.
38. Wohrl S, Vigl K, Zehetmayer S, Hiller R, Jarisch R, Prinz M, et al. The performance of a component-based allergen-microarray in clinical practice. Allergy. 2006; 61:633–9. [PubMed: 16629796]
39. Shin DS, Compadre CM, Maleki SJ, Kopper RA, Sampson H, Huang SK, et al. Biochemical and structural analysis of the IgE binding sites on Ara h1, an abundant and highly allergenic peanut protein. J Biol Chem. 1998; 273:13753–9. [PubMed: 9593717]
40. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. Clin Biochem Rev. 2008; 29:S83–7. [PubMed: 18852864]
41. Nicolaou N, Poorafshar M, Murray C, Simpson A, Winell H, Kerry G, et al. Allergy or tolerance in children sensitized to peanut: prevalence and differentiation using component-resolved diagnostics. J Allergy Clin Immunol. 2010; 125:191–7. [PubMed: 20109746]
42. Johnson K, Keet C, Hamilton R, Wood R. Predictive value of peanut component specific IgE in a clinical population. J Allergy Clin Immunol. 2011; 127:AB71.
43. Jones SM, Pons L, Roberts JL, Scurlock AM, Perry TT, Kulis M, et al. Clinical efficacy and immune regulation with peanut oral immunotherapy. J Allergy Clin Immunol. 2009; 124:292–300. [PubMed: 19577283]
44. Vickery BP, Lin J, Kulis M, Steele PH, Jones SM, Bardina L, et al. Peanut oral immunotherapy (OIT) induces epitope-specific isotype shift from IgE to IgG_4. J Allergy Clin Immunol. 2010; 125:AB20.
45. Savilahti EM, Rantanen V, Lin JS, Karinen S, Saarinen KM, Goldis M, et al. Early recovery from cow's milk allergy is associated with decreasing IgE and increasing IgG_4 binding to cow's milk epitopes. J Allergy Clin Immunol. 2010; 125:1315–21. [PubMed: 20462631]
46. James LK, Shamji MH, Walker SM, Wilson DR, Wachholz PA, Francis JN, et al. Long-term tolerance after allergen immunotherapy is accompanied by selective persistence of blocking antibodies. J Allergy Clin Immunol. 2011; 127:509–16. e1–5. [PubMed: 21281875]

Abbreviations used

DBPCFCs	Double-blind, placebo-controlled food challenges
HSA	Human serum albumin
ROC	Receiver operating characteristics
sIgE	Specific IgE
SPT	Skin prick test

Clinical implications

A novel peanut allergy diagnostic approach with higher accuracy than current allergy tests was developed by employing peptide microarray immunoassays and bioinformatics. This method may be useful for clinical allergy testing in the future.

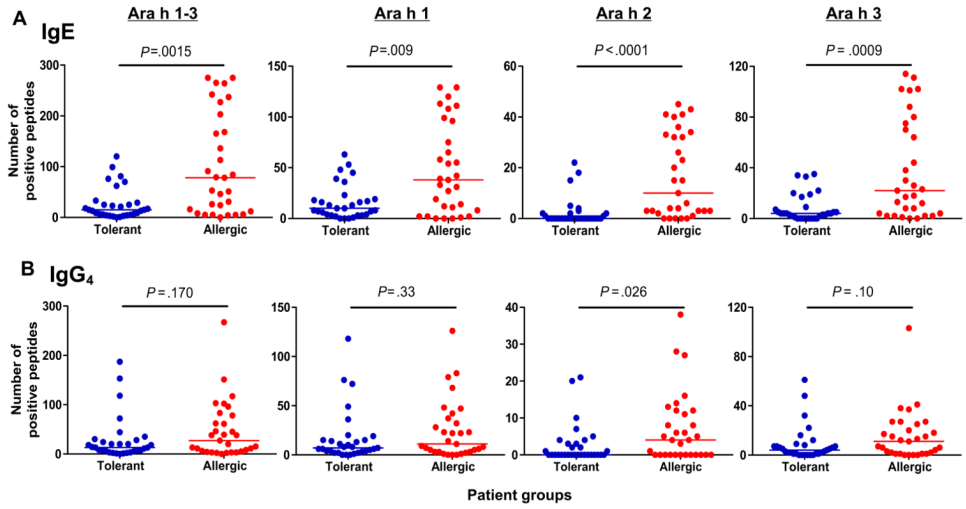

Fig 1.

Comparison of IgE (**A**) and IgG_4(**B**) binding diversity to peanut allergens Ara h 1, Ara h 2, and Ara h 3, and binding diversity to individual allergen Ara h 1, Ara h 2, and Ara h 3 between peanut-allergic and peanut-tolerant groups. Antibody binding diversity was measured as the number of positive peptides (robust Z score > 3) determined by using peptide microarray immunoassay.

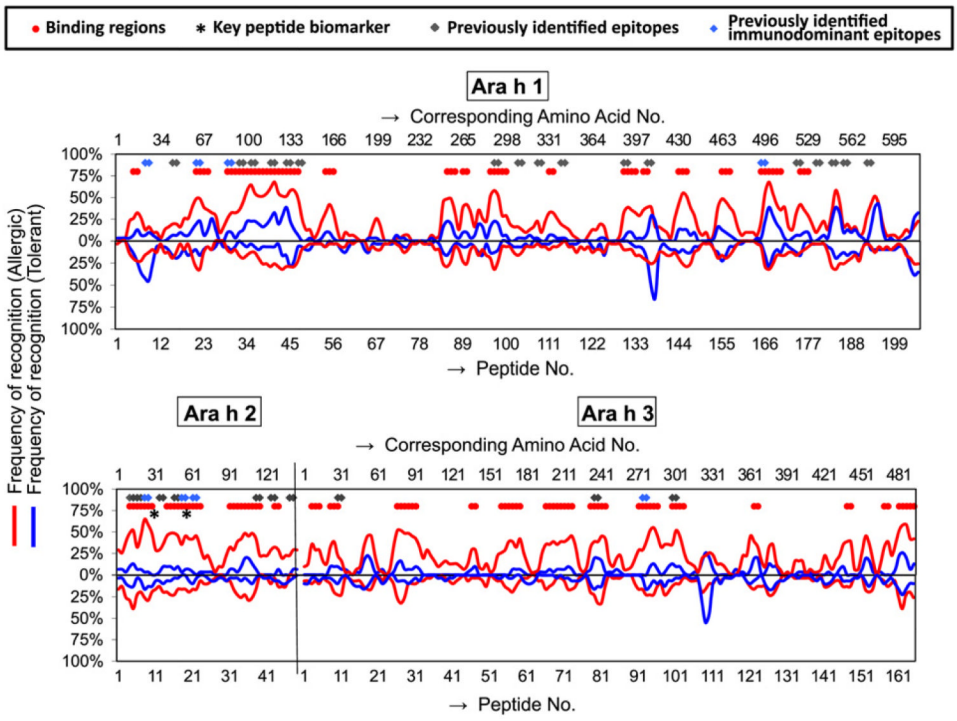

Fig 2.

Comparison of IgE (above x-axis line) and IgG_4 binding (below x-axis line) to peptides of Ara h 1, Ara h 2, and Ara h 3 between peanut-allergic (*red lines*) and peanut-tolerant (*blue lines*) groups. The bottom x-axis shows the overlapping peptides, and the top x-axis shows the corresponding amino acid number of the peptide. The y-axis shows the percentage of patients within each group showing positive binding to each peptide. IgE-binding regions/epitopes identified by using TileMap and the key peptide biomarkers identified by using machine learning methods are indicated with *red circles* and *asterisks*, respectively. The previously identified epitopes and immunodominant epitopes are indicated with *gray* and *blue diamonds*, respectively.

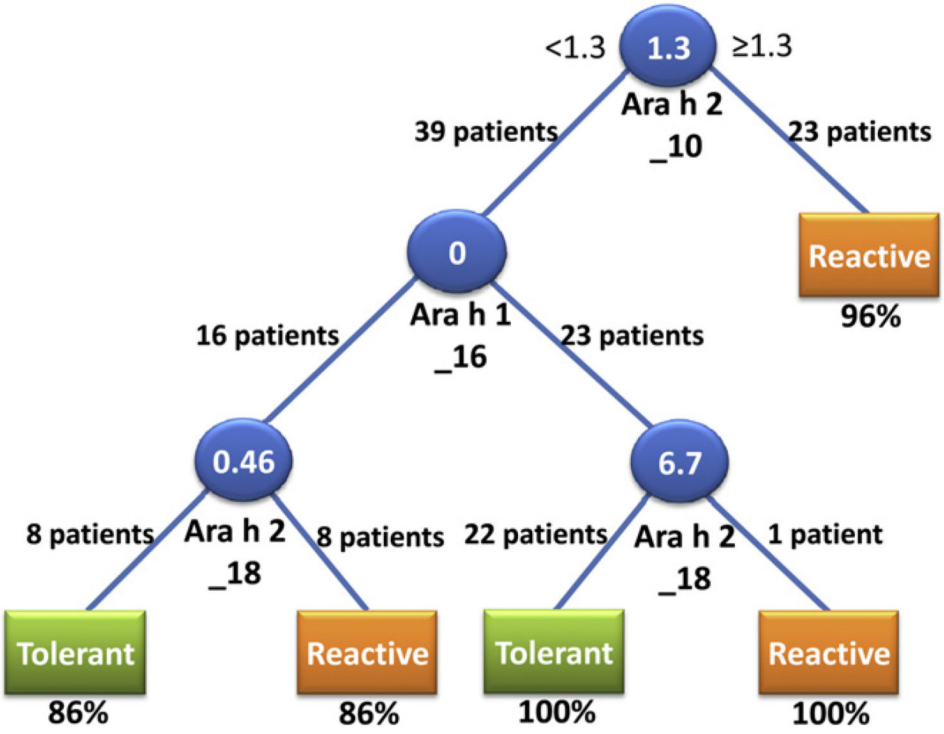

Fig 3.

Decision tree built for classifying peanut-allergic and peanut-tolerant individuals. Sixty-two individuals were sorted from root (*top circle*) to leaf nodes (*rectangles*), based on their IgE reaction to a panel of peanut peptide biomarkers defined at each node (*circle*), which represent the splitting points. The peptides selected for each splitting point are listed under the circles, and the splitting threshold (Z score) appears in the circle. At each splitting point, individuals with IgE reactions to the selected peptide at or above the threshold are assigned to the right branch and below the threshold to the left. The value on the branch shows the number of individuals passing through. The percentage value under the leaf nodes represents the calculated accuracy.

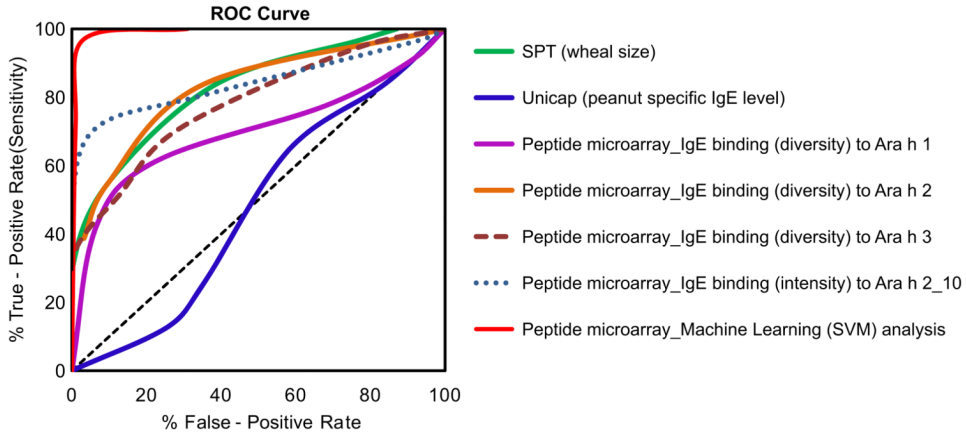

Fig 4.

Comparison of the diagnostic performance of different allergy tests and analysis methods in predicting the outcome of DBPCFC. The area under the ROC curve indicates how well a test method can distinguish between 2 diagnostic groups (peanut allergic vs peanut tolerant). The *diagonal line* indicates a completely random guess. Both IgE binding diversity (expressed as the number of positive peptides) and intensity (express as Z score) were measured by using peptide microarray.

Table I

Patients' demographic and clinical characteristics*

	Peanut-allergic group	Peanut-tolerant group	*P* value
No. of patients	31	31	
Age (y)	7.1 ± 3.7 [3-17]	7.9 ± 3.7 [3-17]	NS
Males (n)	19 (61.1%)	20 (67.7%)	NS
Wheal diameter (mm)	7.2 ± 3.8 [3-13]	3.8 ± 2.1 [0-8]	<.0001
sIgE (kU_A/L)	2 ± 2.7 [0.35-15.1]	2.4 ± 2.9 [0.35-13.4]	NS
Other food allergy	21 (67.7%)	23 (74.1%)	NS
Atopic dermatitis	22 (70.6%)	21 (67.7%)	NS
Asthma	15 (48.3%)	15 (48.3%)	NS
Allergic rhinitis	21 (67.7%)	23 (74.1%)	NS
Urticaria	3 (9.6%)	0 (0%)	NS

NS, Not significant.

*Values are expressed as mean ± SD, median [range], or number (%).

Table II

Performance of decision tree and support vector machine classifier for predicting the patients' allergic status by 5-fold cross-validation*

	Decision tree	Support vector machines
Specificity	0.94 (0.04)	0.97 (0.03)
Sensitivity	0.87 (0.06)	0.90 (0.05)
Overall accuracy Q2†	0.90 (0.04)	0.94 (0.03)

*SEs are shown in parentheses. The overall accuracy calculated here indicates the performance of predicting new peanut-allergic individuals in the test data set.

†Overall accuracy Q2 = p/N, where p is the total number of correctly predicted cases and N is the total number of cases.