

Peanut epitopes for IgE and IgG4 in peanut-sensitized children in relation to severity of peanut allergy

Annebeth E. Flinterman, MD,^a Edward F. Knol, PhD,^a Doerthe A. Lencer, MD,^b Ludmilla Bardina, MS,^b Constance F. den Hartog Jager, BSc,^a Jing Lin, PhD,^b Suzanne G. M. A. Pasmans, MD, PhD,^a Carla A. F. M. Bruijnzeel-Koomen, MD, PhD,^a Hugh A. Sampson, MD,^b Els van Hoffen, PhD,^a and Wayne G. Shreffler, MD, PhD^b *Utrecht, The Netherlands, and New York, NY*

Background: Better understanding of the relationship between antibody response to peanut and clinical sensitivity might lead to more accurate prognostication.

Objective: We sought to investigate peanut-specific IgE and IgG4 epitope diversity in relation to challenge-defined clinical sensitivity to peanut in a group of peanut-sensitized children.

Methods: Clinical sensitivity was determined by means of double-blind, placebo-controlled peanut challenges in 24 sensitized children. Six atopic control subjects were included. Specific IgE and IgG4 binding to 419 overlapping 15-amino-acid peptides representing the sequence of recombinant Ara h 1, Ara h 2, and Ara h 3 was analyzed by means of microarray immunoassay.

Results: Peanut-sensitized patient sera bound significantly more IgE and IgG4 epitopes than control sera. This patient group reacted to the same Ara h 1, Ara h 2, and Ara h 3 epitopes as reported previously. There was a positive correlation between IgE epitope diversity (ie, number of epitopes recognized) and clinical sensitivity ($r = 0.6$), such that patients with the greatest epitope diversity were significantly more sensitive than those with the lowest diversity ($P = .021$). No specific epitopes were associated with severe reactions to peanut. IgG4 binding was observed to largely similar epitopes but was less pronounced than IgE binding and did not relate to the clinical sensitivity to peanut. IgE and IgG4 epitope-recognition patterns were largely stable over a 20-month period.

Conclusion: Clinical sensitivity, as determined by means of double-blind, placebo-controlled peanut challenge, is positively related to a more polyclonal IgE response, which remains stable over time. (*J Allergy Clin Immunol* 2008;121:737-43.)

Key words: Peanut allergy, children, peptides, IgE, IgG4, severity of allergic reactions, diversity

Peanut allergy is among the most prevalent and severe food allergies in westernized countries.^{1,2} Previous studies suggest that

Abbreviations used

DBPCFC: Double-blind, placebo-controlled food challenge

ED: Eliciting dose

HSA: Human serum albumin

MIA: Microarray immunoassay

PBS-T: PBS containing 0.5% Tween 20

SNR: Signal-to-noise ratio

clinical sensitivity to allergens, including peanut, might be related to allergen-specific IgE epitope pattern, diversity, and avidity, all of which likely play a role in IgE function on effector cells, such as mast cells and basophils.³⁻⁷

Numerous peanut allergens (designated Ara h 1-Ara h 8)⁸ have been characterized, and the sequential IgE-binding epitopes of peanut allergens have been previously defined for several of them.⁹⁻¹³ Several studies have shown in both patients with peanut allergy and sensitized animals that the majority of the IgE response is directed to Ara h 1, Ara h 2, Ara h 3, and Ara h 6.^{3,14,15}

The function of allergen-specific IgE might also be modulated by the presence of other specific antibody classes. Increased specific IgG4 levels, in particular, have been associated with suppression of IgE-dependent immediate hypersensitivity reactions, such as in the context of venom immunotherapy, natural high-dose cat exposure, and helminth infection.¹⁶⁻¹⁸ Allergen-specific IgG subclasses have also been shown to suppress IgE-facilitated antigen presentation, suggesting that they might modulate the progression of adaptive immunity with ongoing allergen exposure.¹⁹

We showed previously that IgE diversity to continuous epitopes of Ara h 1 through Ara h 3 correlated with reaction severity, as well as *in vitro* effector cell degranulation.⁴ However, clinical reactions were not defined by double-blind, placebo-controlled peanut challenge (DBPCFC), and the study was limited by its retrospective design. Lewis et al³ also demonstrated a relationship between IgE diversity, as defined by immunoblotting, and clinical sensitivity.

The aim of the present study was to prospectively determine the relationship between peanut allergen epitope-specific IgE and IgG4 diversity and clinical sensitivity defined by DBPCFC.

METHODS

Study population

Twenty-four children (8 female and 16 male children; age range, 3-15 years; mean age, 7.2 years) were recruited from the Department of Pediatric Dermatology/Allergy at the University Medical Centre Utrecht. These children participated in a previous study to determine eliciting doses (EDs) for peanut and were a representative sample from the population of peanut-

From ^athe Department of Dermatology/Allergy, University Medical Centre Utrecht, and ^bthe Department of Pediatrics, Division of Allergy and Immunology, and the Immunology Institute, Mount Sinai School of Medicine, New York.

Supported in part by a research grant obtained from the Wilhelmina Children's Research Foundation, The Netherlands. W.G.S. was supported by NIH K12HD052890 and the Pediatric Research NIH-LRP.

Disclosure of potential conflict of interest: L. Bardina has received research support from the National Institutes of Health and the Food Allergy Research Initiative. H. A. Sampson has consulting arrangements with and owns stock in Allertein. The rest of the authors have declared that they have no conflict of interest.

Received for publication July 15, 2007; revisions received November 27, 2007; accepted for publication November 28, 2007.

Available online January 31, 2008.

Reprint requests: Edward F. Knol, PhD, University Medical Centre, G02.124, PO Box 85500, 3508 GA Utrecht, The Netherlands. E-mail: e.f.knol@umcutrecht.nl.

0091-6749/\$34.00

© 2008 American Academy of Allergy, Asthma & Immunology

doi:10.1016/j.jaci.2007.11.039

sensitized children in our outpatient clinic with regard to specific IgE levels and severity of previous allergic reactions to peanut.²⁰ Inclusion criteria consisted of increased peanut-specific IgE levels. Six atopic but not peanut-sensitized children were included as control subjects. All parents provided written informed consent before enrollment in the study. The study was reviewed and approved by the Central Committee of Human-based Research in The Netherlands (CCMO, The Hague, The Netherlands).

DBPCFCs

DBPCFCs were performed as described previously.²⁰ In short, increasing amounts of defatted peanut flour (0.01, 0.1, 0.3, 0.5, 1, 10, 100, 300, 1000, and 3000 mg) or placebo were given. Although all challenges were continued until there were objective symptoms, the first dose to provoke symptoms, either subjective or objective, was used for determination of the ED. The severity of allergic reactions was scored according to the method of Mueller²¹ as mild (grade 0 and 1), moderate (grade 2), or severe (grade 3 and 4). Mild reactions were symptoms of the oral cavity, the skin, and/or the mucous membranes; moderate reactions included gastrointestinal symptoms; and severe reactions consisted of respiratory symptoms and/or cardiovascular symptoms.

Peanut-specific IgE levels

Serum samples were collected for each patient before and 20 months after DBPCFCs. Peanut-specific IgE levels were determined in all samples by using the CAP System FEIA, according to the manufacturer's instructions (Phadia, Uppsala, Sweden). Specific IgE levels of greater than 0.35 kU/L were considered positive.

Peptides, reagents, and array production

We determined IgE and IgG4 binding to sequential epitopes of Ara h 1, Ara h 2, and Ara h 3 using a peptide microarray-based immunoassay.^{4,13} A library of peptides, 15 amino acids in length and with an offset of 3, corresponding to Ara h 1, Ara h 2, and Ara h 3, was commercially synthesized (JPT Technologies GmbH, Berlin, Germany). Peptides included a 4-residue linker sequence and a reactive group for site-specific binding.

Peptides were resuspended at 150 μ mol/L in Protein Printing Buffer (ArrayIt, Sunnyvale, Calif) and printed in 2 sets of duplicates onto epoxy-derivatized glass slides (SuperEpoxy Substrate, ArrayIt) by using the NanoPrint Microarrayer 60 (TeleChem International, Inc, Sunnyvale, Calif). Slides were stored at 4°C until use. Additional array elements, including purified Ara h 1 and whole peanut extract, were used as positive controls. Protein Printing Buffer alone was used as a negative control and for background normalization. Fluorochrome-labeled BSA elements were used for grid alignment (positional control).

Immunolabeling

The slides were rinsed with PBS containing 0.5% Tween 20 (PBS-T), and nonspecific binding sites were blocked with 1% human serum albumin (HSA) in PBS-T for 1 hour (PBS-T/HSA). After removing the PBS-T/HSA from the slide surface, 80 μ L of patient sera diluted 1:6 in PBS-T/HSA was incubated for 1 hour on a rotator. Polyclonal goat anti-human IgE diluted 1:5000 (a gift from Phadia, Uppsala, Sweden) combined with monoclonal mouse anti-human IgG4 diluted 1:10,000 (clone G174; PharMingen, San Jose, Calif), which were covalently tagged with Alexa 546 and Alexa 647 (Invitrogen, Carlsbad, Calif), respectively, were then incubated for 1 hour. All incubations were performed in a humidity chamber (Binding Site, Birmingham, United Kingdom) at room temperature in the dark. Slides were washed with PBS-T, centrifuged dry, and scanned with a ScanArray Gx (Perkin Elmer, Waltham, Mass). Images were saved as TIF files. Fluorescence signal was digitized with the program ScanArray Express (Perkin Elmer).

Analysis

Data were exported as comma-delimited text files and analyzed by using R analysis.²² The complete raw data set and a script for all analyses and

generation of all figures is available in Table E1 in this article's Online Repository at www.jacionline.org. The readout used for both IgE and IgG4 binding was the median fluorescence signal-to-noise ratio (SNR) of the 4 replicates for each array element. Briefly, median SNR values were log² transformed and median centered for interslide comparison. Array features with an intraslide coefficient of variation of greater than 5% were excluded from further analysis (<1% of total). Based on our previous data¹³ and the published literature on B-cell epitope length, we defined positive binding as necessarily involving at least 2 overlapping peptides (maximum length, 12 residues), and therefore subsequent analysis was carried out on data for each element adjusted to the median of itself and the 2 flanking peptides (see Fig E1 in the Online Repository at www.jacionline.org for display of data without this near-neighbor normalization). Fig E2 (available in the Online Repository at www.jacionline.org) shows histograms for IgE (Fig E2, A) and IgG4 (Fig E2, B) data derived from control sera. For calculation of percent positive values for each peptide, cutoff values for IgE and IgG4 were determined from these control data at a *P* value of .995 by using the following formula: $p(k) = (k - 1/3) / (n + 1/3)$. The resulting probability estimates are approximately median unbiased, regardless of the distribution of the data.

We defined the clinical sensitivity as the product of the Mueller symptom score and the ED score (Table I), and this was treated as a ranked variable. For correlation of clinical sensitivity, symptom score, or ED score with epitope number, the Spearman rank correlation test was used. For comparison of clinical sensitivity score between the highest and lowest epitope quartile groups, the Wilcoxon test was used. The χ^2 test was used to assess the significance of agreement of IgE and IgG4 binding patterns over time.

RESULTS

IgE and IgG4 reactivity to peptides in healthy control subjects and study subjects

IgE and IgG4 binding to peptides from Ara h 1, Ara h 2, and Ara h 3 was detected simultaneously by using 2 distinct fluorescence channels. For the nonallergic control subjects, IgE and IgG4 fluorescence was low, and by definition, signal of greater than the arbitrary cutoff was low (<0.5% for both; Fig 1, A; see the Methods section for cutoff determination). By using the same cutoffs for the patient group, 3.1% for IgE (312 positive peptides of a total of 9967 peptides studied in the patient group) and 1.8% for IgG4 (182 of 9967 peptides) had signal of greater than the cutoff levels for each channel alone, and an additional 0.8% (77/9967) were positive for both IgE and IgG4 (Fig 1, B).

The number of positive IgE epitopes per individual ranged from 0 to 80 (median, 9) and correlated with the level of peanut-specific IgE ($r = 0.579$, see Fig E3 in the Online Repository at www.jacionline.org). Overall, 22 (92%) of 24 patients recognized at least 1 epitope from the 3 allergens. IgE and IgG4 binding was clustered to epitope-rich areas (Fig 2, A). There were 17 distinct IgE signal maxima in Ara h 1, 5 in Ara h 2, and 5 in Ara h 3. One area in Ara h 3 not previously investigated by means of microarray assay⁴ was positive across a range of peptides, with only 3 overlapping amino acids, IYR, for IgE and IgG4 in both the study and control subjects, and this is a suspected artifact of unknown significance. The IYR sequence was not present in other peptides. Excluding this sequence, the most dominant epitopes (Ara h 1: PSHQQPRKI and FYFPSRRFS; Ara h 2: RRCQSQ; and Ara h 3: EDEYEDDEE) were recognized by 30% to 40% of patients, which is consistent with previous results.⁴

Clinical evaluation of peanut allergy

The severity of the symptoms and the EDs were determined by means of DBPCFC, as previously described, to relate the IgE and IgG4 binding epitopes to the clinical reactivity to peanut.²⁰

TABLE I. Age, sex, specific IgE to peanut, and results of DBPCFC per subject, followed by the calculation of the combination score as symptom score $\times$ ED score

Patient ID	Age (y)	Sex	IgE*	Symptoms	Score	ED†	Score	Combination	IgE epitopes	IgG4 epitopes
Pi5	3	Male	4	—	0	—	0	$0 \times 0 = 0$	3	0
Pi9	6	Male	7	Rc, vom	2	300	2	$2 \times 2 = 4$	0	24
Pi11	5	Female	250	Rc, hoa	2	100	5	$2 \times 5 = 10$	8	1
Pi12	10	Female	50	—	0	—	0	$0 \times 0 = 0$	0	0
Pi14	6	Male	0.4	Oas	1	300	2	$1 \times 2 = 2$	3	13
Pi15	14	Male	2.7	Oas, rc, ap	2	>30	1	$2 \times 1 = 2$	0	3
Pi16	15	Female	35	Oas, rc, vom	2	300	2	$2 \times 2 = 4$	8	50
Pi17	6	Male	527	Oas, vom	2	10	6	$2 \times 6 = 12$	61	64
Pi19	6	Female	183	Rc, vom, bro	3	100	3	$3 \times 3 = 9$	18	7
Pi20	5	Female	4	Oas, rc, ap, dia, hoa	2	300	2	$2 \times 2 = 4$	4	15
Pi21	12	Male	20	Oas, rc, ap, vom	2	300	4	$2 \times 4 = 8$	4	3
Pi22	10	Female	61	Oas, ap, vom, dia	2	300	4	$2 \times 4 = 8$	27	8
Pi24	4	Male	815	Gu, vom, str	3	100	3	$3 \times 3 = 9$	80	13
Pi25	4	Male	1	Oas, fs, ap	2	100	3	$2 \times 3 = 6$	19	13
Pi26	5	Female	48	Gu, rc, vom	2	300	2	$2 \times 2 = 4$	10	6
Pi27	3	Male	8	Oas, fs, rc, ap, hoa	2	100	3	$2 \times 3 = 6$	4	8
Pi28	5	Male	99	Oas, gu, rc, bro, tach	3	100	3	$3 \times 3 = 9$	11	5
Pi29	4	Female	15	Oas, ap, vom	2	100	3	$2 \times 3 = 6$	10	1
Pi36	7	Male	163	Oas, ap	2	10	6	$2 \times 6 = 12$	39	4
Pi37	12	Male	3	Oas, fs, urt, rc, ap	2	100	5	$2 \times 5 = 10$	8	0
Pi38	7	Male	93	Gu, rc, ap, vom, dia	2	>30	1	$2 \times 1 = 2$	41	6
Pi39	4	Male	2	Oas, vom, rc	2	300	4	$2 \times 4 = 8$	4	0
Pi40	12	Male	3	Nau, vom	2	100	3	$2 \times 3 = 6$	16	9
Pi41	7	Male	35	Oas, cu, fs	1	300	2	$1 \times 2 = 2$	11	6

rc, Rhinoconjunctivitis; vom, vomiting; hoa, hoarseness; Oas, oral allergy syndrome; ap, abdominal pain; bro, bronchoconstriction; dia, diarrhea; gu, generalized urticaria; str, stridor; fs, facial swelling; tach, tachycardia; urt, urticaria; cu, contact urticaria.

*Specific IgE in kilounits per liter.

†ED is defined as eliciting dose in milligrams of peanut flour.

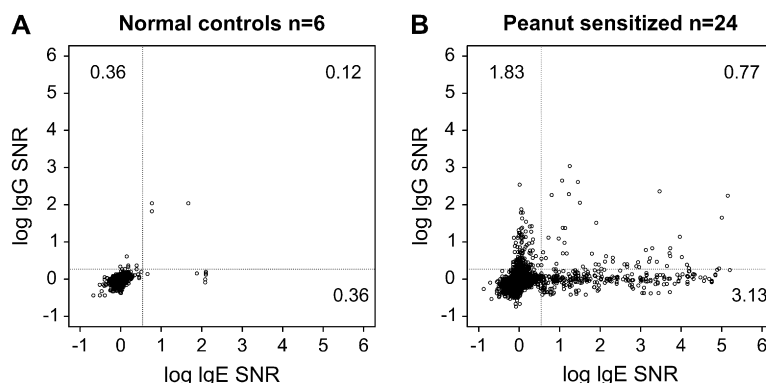


FIG 1. IgE and IgG4 peptide binding to peanut allergen–derived peptides is significantly greater in sera from peanut-sensitized patients versus sera from control subjects. IgE SNR versus IgG4 SNR is shown. Dashed lines indicate cutoffs (see the Methods section). Numbers indicate percentages above cutoff. **A**, Control sera ($n = 6$). **B**, Peanut-sensitized sera at baseline ($n = 24$).

Challenge outcomes are displayed together with demographic data and peanut-specific IgE levels in Table I. Twenty-two (92%) children had an allergic reaction to peanut during DBPCFC, ranging from isolated oral pruritis to multisystemic reactions with severe bronchospasm, generalized urticaria, and tachycardia.

The individual EDs ranged from 10 mg to greater than 3 g. To take both parameters into account, we used a scoring for the clinical sensitivity according to a combination of severity of symptoms (mild, moderate, and severe) and EDs (step 1-6, corresponding to reactions from >3000 mg to reactions from

10 mg). The symptom score was multiplied by the ED score to give a “combination score” conceptually similar to Lewis et al³ (Table I).

IgE reactivity to peptides in relation to the clinical sensitivity to peanut

Specific IgE-binding regions for children with either low or high combination scores were not found. Individual epitope-binding patterns were variable and unique (Fig 2, B).

A positive relationship between the clinical sensitivity to peanut, defined by the combination score, and the number of

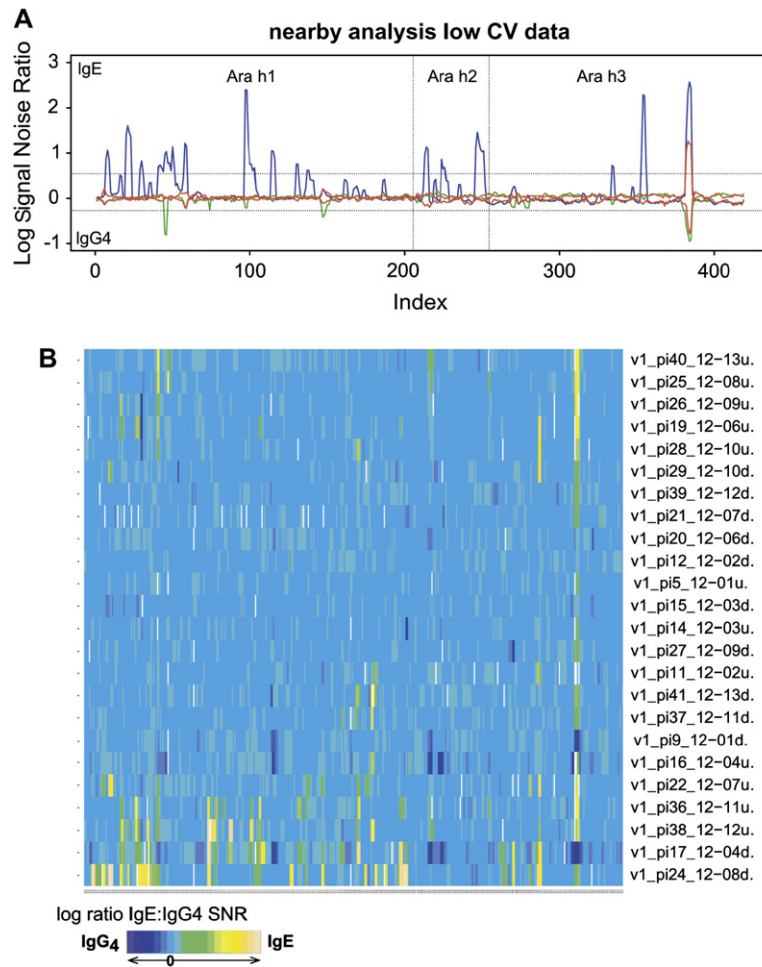


FIG 2. Multiple sequential epitope-rich regions within the peanut allergens Ara h 1 through Ara h 3. **A**, Representation of mean log SNR for binding of patient IgE (blue line, $n = 24$) and IgG4 (green line, $n = 24$) versus control (red lines, $n = 6$). Dashed lines indicate cutoff values (see the Methods section and Fig E2). **B**, Heat map image of individual IgE/IgG4 SNR showing diversity of individual binding patterns (see the Methods section for details).

positive IgE epitopes was found (Spearman $r = 0.58$). When subjects were binned into quartiles based on the number of recognized epitopes, a trend was observed for higher clinical sensitivity with greater IgE epitope diversity, which reached statistical significance for the comparison between the highest and lowest quartiles (Fig 3, A; median combination score of 8.5 [range, 2-12] vs 4 [range, 0-8] for groups 4 and 1, respectively; Wilcoxon $P = .0214$).

IgE and IgG4 reactivity were also compared after grouping by combination score (Fig 4). The quartile of patients with the highest clinical sensitivity had positive IgE binding to approximately 7.4% of all peptides compared with 1.9%, 1.2%, and 2.0% for the first-, second-, and third-quartile groups, respectively. Two children appeared not clinically allergic on challenge, including one with a level of specific IgE previously determined to be highly suggestive of clinical reactivity (Pi12 in Table I; 50 kU/L, Table I).²³ Both of these individuals had very little epitope binding, as determined by means of microarray assay. In the case of Pi12, there were no positive epitopes, and for Pi5, only Ara h 3 IYR was positive. Analysis excluding these 2 nonallergic individuals did not change the observed

relationship between IgE epitope diversity and higher combination scores.

The relationship between IgE binding and the combination score was stronger than that of IgE binding to either the symptom score alone ($r = 0.49$) or the ED score alone ($r = 0.32$). IgE epitope number correlated more strongly with clinical sensitivity than specific IgE ($r = 0.58$ vs $r = 0.52$).

IgG4 reactivity in relation to clinical sensitivity

The number of IgG4 epitopes varied from 0 to 64 per subject (median, 6). IgG4 epitopes were mainly located on Ara h 1 and Ara h 3, with epitope densities of 2.2% (percentage of peptides with SNR of greater than the cutoff level) and 3.5%, respectively. In contrast, Ara h 2 had a density of IgG4 epitopes of 1.0%. No relationship was observed between IgG4 diversity and clinical sensitivity.

IgG4 binding coincided with IgE binding in 29.6% of the IgG4-positive peptides (Fig 1). Consistent with previous analysis of Ara h 2,¹³ this was more than expected on the basis of the frequency of IgE and IgG4 alone (0.77% double-positive results observed vs 0.1% expected; $P < .001$, χ^2 test).

IgE and IgG4 reactivity after 20 months

Previously, we demonstrated that the IgE binding to the major peanut allergens by means of Western immunoblotting was not changed at 20 months after DBPCFC in these subjects.²⁴ Microarray immunoassays (MIAs) were repeated with serum samples obtained 20 months after the challenge to determine the stability of the epitope diversity over time. At that time point, from 2 subjects, no serum samples were available. For the remaining 22 children, in general, epitope diversity was highly similar ($P < .001$ for frequency of agreement between baseline and 20 months for both IgE and IgG4, χ^2 test; Fig 5). IgE binding patterns were more stable over time than IgG4 patterns. We did not observe a substantial change in IgE epitope diversity in any of the subjects 20 months after challenge.

DISCUSSION

The goal of the present study was to investigate in children with peanut allergy the IgE and IgG4 peanut allergen epitope recognition in relation to clinical reactivity, as determined by means of DBPCFC. We found that children with more severe reactions recognized more diverse IgE epitopes, suggesting a more polyclonal response, whereas IgG4 epitope recognition was less abundant and did not relate to clinical sensitivity to peanut. Binding was clustered to previously described epitope-rich areas. However, as we observed before using this method, each individual displayed a unique fingerprint of IgE and IgG4 binding, and these individual binding patterns remained largely stable 20 months later.

The peptide MIA used in this study has been documented before as a reliable tool to describe sequential IgE epitopes on peanut allergens.^{4,13} The reproducibility of this method was demonstrated here by the finding that IgE-binding epitopes on Ara h 1, Ara h 2, and Ara h 3 were similar to the epitopes described in other populations with peanut allergy by using different methods^{9,11,25} and the reproducibility of epitope recognition in these patients over time.

The majority of IgE epitopes were located on Ara h 1 and Ara h 2. The dominant IgE binding to Ara h 2-derived peptides was in line with previous findings with Western immunoblotting, showing that Ara h 2 is an important peanut allergen.^{3,24} In the present population multiple IgE-binding epitope regions on Ara h 1 were defined as well. This is consistent with previous mapping data but contrasts somewhat with results obtained by means of skin prick testing and Western immunoblotting in the current population of patients with peanut allergy.²⁴ Differences between the sequential structure of Ara h 1 in the current study compared with the tertiary structure of purified Ara h 1 might be an explanation. Another difference between purified Ara h 1 and the peptides used in the current study, which were derived from a recombinant Ara h 1, is that a fragment of 86 amino acids on the N-terminus is cleaved off during synthesis. Several epitopes are located on this fragment.¹² Excluding these epitopes demonstrated that the results from MIAs and Western immunoblots were in agreement (not shown). It has been reported that proteolytic fragments after enzymatic digestion of Ara h 1 retained IgE-binding epitopes.²⁶ This might suggest the relevance of sensitization to both Ara h 1 and Ara h 2 in subjects with peanut allergy, which has been implied before for children with peanut allergy.^{15,27}

A significant strength of this study over previous investigations of IgE epitopes in relation to clinical sensitivity is the fact that

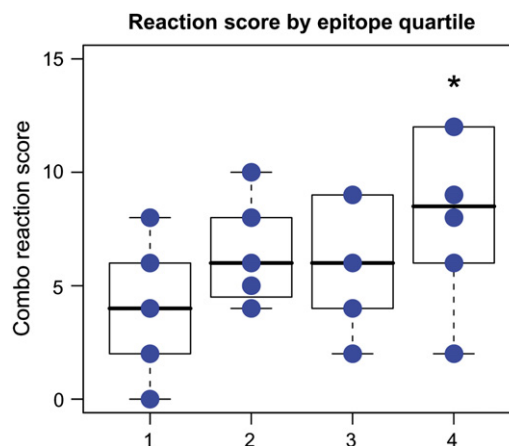


FIG 3. Greater IgE epitope recognition corresponds to greater clinical sensitivity. The distribution of reaction scores is shown for each quartile of patients binned into quartiles according to the number of positive IgE epitopes. Boxes represent median and 25% to 75% distribution, whiskers represent range, and solid circles represent individual reaction scores. Patients in the highest quartile for IgE epitope number (quartile 4, $n = 6$), had significantly higher combination scores in comparison with those with the least IgE binding (quartile 1, $n = 6$). $*P = .021$.

all subjects were characterized prospectively by means of DBPCFC. To define clinical sensitivity, we combined the severity of symptoms and the ED. As Hourihane and colleagues have suggested, the ED is a clinically relevant parameter in combination with the severity of the allergic symptoms.^{3,28} Interestingly, IgE epitope diversity correlated better with the combination score than with either symptom score or the ED score alone.

The diversity of IgE binding was unique for each subject, and IgE epitopes were recognized by a maximum of 30% to 40% of subjects. Therefore no specific immunodominant regions could be related to more severe reactions. Previous studies have also suggested that the number of allergens recognized by IgE rather than the recognition of particular allergens was a risk factor for more serious reactions.^{3,29} Our findings underline that the polyclonality of the antibody response is related to the clinical sensitivity to peanut. Presentation of clustered epitopes might facilitate IgE cross-linking and basophil degranulation.⁷ In support of this, we previously showed that serum containing IgE to a larger number of IgE epitopes on peanut decreased the threshold for basophil and mast cell degranulation.⁴ The diversification of epitope recognition through a process of epitope spreading is a regulated T cell-dependent process that proceeds in concert with affinity maturation.^{30,31}

In contrast to the relatively high frequency of IgE epitopes recognized in these patients, IgG4 binding was low. We hypothesized that IgG4 epitope number might inversely correlate with clinical sensitivity because IgG4 has been shown in other contexts to inhibit the activity of IgE by competing for binding sites.¹⁹ However, no association between IgG4 binding and clinical sensitivity was demonstrated. Somewhat surprising to us was the finding that IgG4 reactivity, even in the 2 peanut-sensitized subjects without peanut allergy, was nearly absent. These low levels of IgG4 might be a reflection of very low exposure to peanut during elimination diets before the DBPCFC because in the contexts in which IgG4 has been correlated with protection, such as cat allergy, it is associated with chronic antigen exposure.¹⁷ The relevance of IgG4 in peanut allergy needs to be further investigated

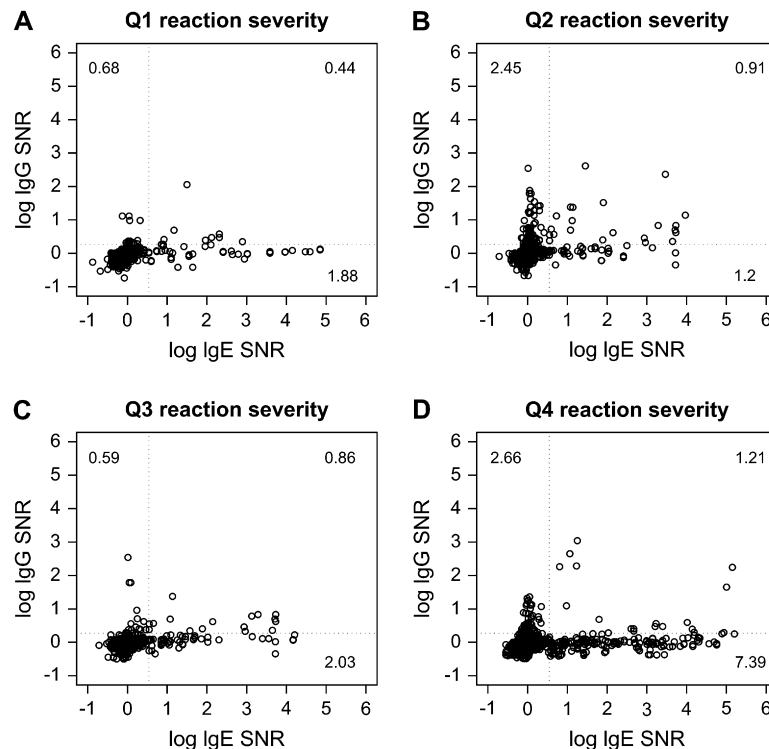


FIG 4. Clinical sensitivity corresponds to greater IgE epitope recognition. IgE versus IgG SNR is shown in separate plots by quartiles of clinical sensitivity (**A** to **D** for least to most sensitive; see the Methods section). Dashed lines represent cutoffs. Numbers in each quadrant represent percentage above cutoff.

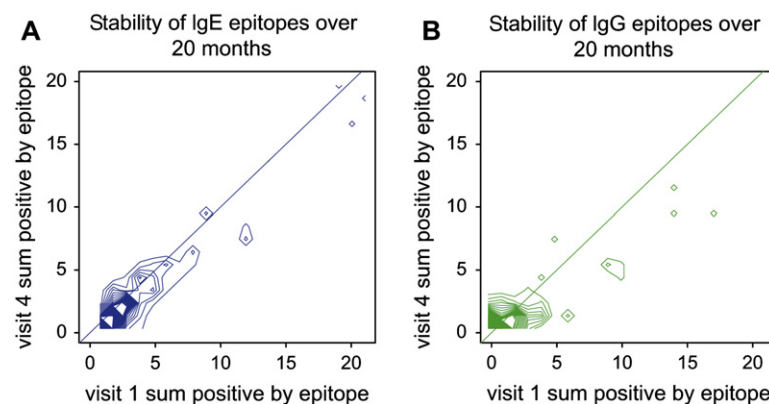


FIG 5. IgE and IgG4 binding patterns are largely stable over 20 months. Correspondence between baseline and 20-month follow-up is shown as a topographic representation of the number of patients with positive results for each peptide before (*x-axis*) and 20 months after (*y-axis*) DBPCFC (*n* = 22). **A**, IgE; **B**, IgG4. The diagonal represents identity between *t*=0 and *t*=20 months.

and compared with other food allergens, such as cow's milk, for which we did observe high levels of IgG4 reactivity with ELISA³² and with MIAs (unpublished observations).

The peanut-specific IgE and IgG4 binding patterns remained largely intact 20 months later. In this period exposure to peanut occurred during DBPCFCs, and possibly there was exposure to trace amounts in the subsequent elimination period. Five children reported inadvertent ingestion of peanut in that period, without significant changes in the IgE or IgG4 profile (data not shown). Exposure to allergens has been suggested to mature the B-cell response, with expanded serologic and clinical reactivity.³³ The

observation in this study indicates that the diversity of IgE and IgG4 binding to peanut was present before DBPCFC and did not develop into an enhanced polyclonal response after a 20-month period.

In conclusion, this study shows that the B-cell response in children with more severe allergic reactions to peanut was characterized by a more diverse peanut-specific IgE response directed mainly to Ara h 1 and Ara h 2. Although the relationship between DBPCFC outcomes and accidental reactions is not simple and additional prospective studies are necessary, these findings suggest that determination of IgE epitope binding

patterns could become the basis of a prognostic test in children with peanut allergy used to identify those at highest risk for severe reactions.

We thank Dr Maarten Hoekstra and Dr André Knulst for their help in recruitment of subjects and critical reading of the manuscript and Todd Martinsky for technical advice.

Clinical implications: Children with more severe allergic reactions to peanut can be characterized by greater diversity of IgE epitopes.

REFERENCES

1. Sicherer SH, Furlong TJ, Munoz-Furlong A, Burks AW, Sampson HA. A voluntary registry for peanut and tree nut allergy: characteristics of the first 5149 registrants. *J Allergy Clin Immunol* 2001;108:128-32.
2. Sicherer SH, Sampson HA. 9. Food allergy. *J Allergy Clin Immunol* 2006;117(suppl):S470-5.
3. Lewis SA, Grimshaw KE, Warner JO, Hourihane JO. The promiscuity of immunoglobulin E binding to peanut allergens, as determined by Western blotting, correlates with the severity of clinical symptoms. *Clin Exp Allergy* 2005;35:767-73.
4. Shreffler WG, Beyer K, Chu T-HT, 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.
5. Mitchell AJ, Moss ND, Collins AM. The biological activity of serum IgE changes over the course of a primary response. *Scand J Immunol* 2002;55:33-43.
6. Boyano-Martinez T, Garcia-Ara C, Diaz-Pena JM, Martin-Esteban M. Prediction of tolerance on the basis of quantification of egg white-specific IgE antibodies in children with egg allergy. *J Allergy Clin Immunol* 2002;110:304-9.
7. Knol EF. Requirements for effective IgE cross-linking on mast cells and basophils. *Mol Nutr Food Res* 2006;50:620-4.
8. SDAP: structural database of allergenic proteins. Available at: http://fermi.utmb.edu/SDAP/sdapf_src.html. Accessed January 17, 2008.
9. 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.
10. 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.
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.
12. 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.
13. Shreffler WG, Lencer DA, Bardina L, Sampson HA. IgE and IgG4 epitope mapping by microarray immunoassay reveals the diversity of immune response to the peanut allergen, Ara h 2. *J Allergy Clin Immunol* 2005;116:893-9.
14. van Wijk F, Hartgring S, Koppelman SJ, Pieters R, Knippels LM. 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.
15. 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.
16. Michils A, Baldassarre S, Ledent C, Mairesse M, Gossart B, Duchateau J. Early effect of ultrarush venom immunotherapy on the IgG antibody response. *Allergy* 2000;55:455-62.
17. Platts-Mills TA, Vaughan JW, Blumenthal K, Pollart Squillace S, Sporik RB. Serum IgG and IgG4 antibodies to Fel d 1 among children exposed to 20 microg Fel d 1 at home: relevance of a nonallergic modified Th2 response. *Int Arch Allergy Immunol* 2001;124:126-9.
18. Hussain R, Poindexter RW, Ottesen EA. Control of allergic reactivity in human filariasis. Predominant localization of blocking antibody to the IgG4 subclass. *J Immunol* 1992;148:2731-7.
19. van Neerven RJ, Knol EF, Ejmaes A, Wurtzen PA. IgE-mediated allergen presentation and blocking antibodies: regulation of T-cell activation in allergy. *Int Arch Allergy Immunol* 2006;141:119-29.
20. Flinterman AE, Pasmans SG, Hoekstra MO, Meijer Y, van Hoffen E, Knol EF, et al. Determination of no-observed-adverse-effect levels and eliciting doses in a representative group of peanut-sensitized children. *J Allergy Clin Immunol* 2006;117:448-54.
21. Mueller HL. Diagnosis and treatment of insect sensitivity. *J Asthma Res* 1966;3:331-3.
22. R Development Core Team. R: a language and environment for statistical computing. Version 2.5.0. Vienna, Austria: R Foundation for Statistical Computing; 2007. Available at: <http://www.r-project.org>. Accessed January 17, 2008.
23. Sampson HA. Utility of food-specific IgE concentrations in predicting symptomatic food allergy. *J Allergy Clin Immunol* 2001;107:891-6.
24. Flinterman AE, van Hoffen E, den Hartog Jager CF, Koppelman SJ, Pasmans SG, Hoekstra MO, et al. Children with peanut allergy recognize predominantly Ara h 2 and Ara h 6, which remains stable over time. *Clin Exp Allergy* 2007;37:1221-8.
25. 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.
26. Maleki SJ, Kopper RA, Shin DS, Park CW, Compadre CM, Sampson H, et al. Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation. *J Immunol* 2000;164:5844-9.
27. Bernard H, Paty E, Mondoulet L, Burks AW, Bannon GA, Wal JM, et al. Serological characteristics of peanut allergy in children. *Allergy* 2003;58:1285-92.
28. Hourihane JO, Grimshaw KE, Lewis SA, Briggs RA, Trewin JB, King RM, et al. Does severity of low-dose, double-blind, placebo-controlled food challenges reflect severity of allergic reactions to peanut in the community? *Clin Exp Allergy* 2005;35:1227-33.
29. Peeters KA, Koppelman SJ, van Hoffen E, van der Tas CW, den Hartog Jager CF, Penninks AH, et al. Does skin prick test reactivity to purified allergens correlate with clinical severity of peanut allergy? *Clin Exp Allergy* 2007;37:108-15.
30. Vanderlugt CL, Begolka WS, Neville KL, Katz-Levy Y, Howard LM, Eagar TN, et al. The functional significance of epitope spreading and its regulation by co-stimulatory molecules. *Immunol Rev* 1998;164:63-72.
31. Neuberger MS, Ehrenstein MR, Rada C, Sale J, Batista FD, Williams G, et al. Memory in the B-cell compartment: antibody affinity maturation. *Philos Trans R Soc Lond B Biol Sci* 2000;355:357-60.
32. Ruiter B, Knol EF, van Neerven RJJ, Garssen J, Knulst AC, Bruijnzeel-Koomen CAFM, van Hoffen E. Maintenance of tolerance to cow's milk in atopic individuals is characterized by high levels of specific IgG4. *Clin Exp Allergy* 2007;37:1103-10.
33. Mudde GC, Bheekha R, Bruijnzeel-Koomen CA. Consequences of IgE/CD23-mediated antigen presentation in allergy. *Immunol Today* 1995;16:380-3.

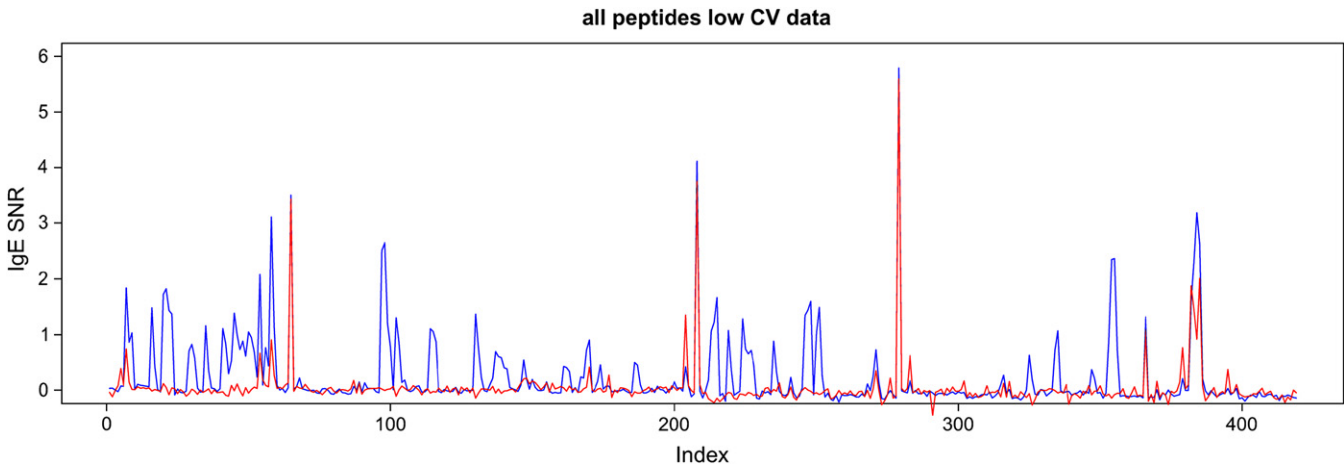


FIG E1. Representation of mean log IgE SNR of raw data by array element before near-neighbor normalization to remove false-positive signal. Patient IgE is shown by the *blue line* (n = 24), and control is shown by the *red line* (n = 6). See Fig 2 for comparison with data after near-neighbor correction. CV, Coefficient of variation.

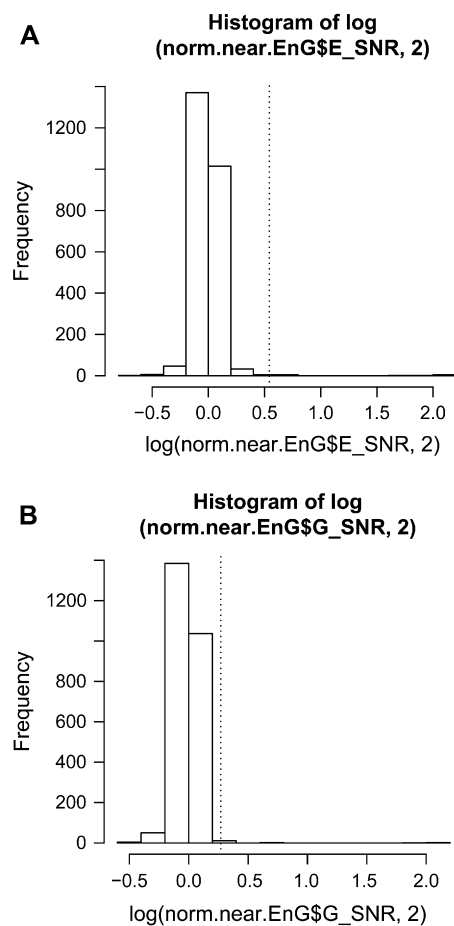


FIG E2. Distribution of signal from control sera ($n = 6$) and determination of cutoff. Histograms of IgE (**A**) and IgG4 (**B**) SNR for peptide array elements incubated with control sera from patients without peanut allergy are shown. The vertical broken line represents the cutoff point at a P value of .995 by using the formula $p(k) = (k-1/3)/(n+1/3)$, which is used for determination of “positive” signal. See Fig 1 for comparison of signal from control and patient sera.

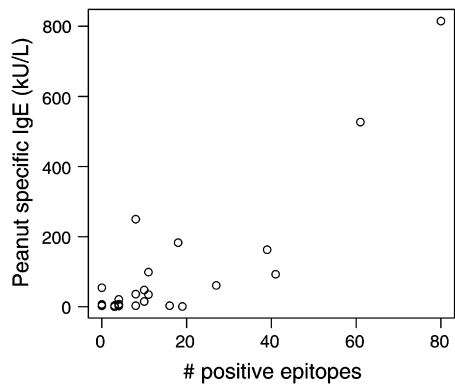


FIG E3. Correlation between peanut-specific IgE levels and the number of epitopes recognized by IgE per subject ($r = 0.579$).

TABLE E1. The complete raw data set and script for all analyses from the study

```
# Copyright 2007, Wayne G. Shreffler. Free non-commercial use is welcome.
# This is a R analysis script for analysis of data set produced in collaboration with Annebeth Flinterman, Edward Knol and others.
# R is by the R Foundation for Statistical Computing and is freely available at http://www.R-project.org. R Development Core Team (2006). R: A language and
  environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL.

# THIS WILL RUN WITHOUT MODIFICATION ON OS X but will require editing of pipe functions and substitution of quartz functions (e.g. with 'pdf') for
  use on Windows OS.

# Run by copying text from this document and pasting to R console.

# Direct questions or suggestions about this script to the author at wayne.shreffler@mssm.edu.

# begin script
#functions
stdev <- function(x) (sqrt(var(x)/length(x)))/mean(x)
#cutoff <- function(x) (mean(x) + 3*stdev(10+x))
current_dir <- readLines(pipe("pwd"))
download.file("http://www.iisina.org/shreffler/flinterman_data.zip", "flinterman_data.zip")
pipe("unzip 'flinterman_data.zip'", "w")

BATCH <-function(directory, exportname="R_sum.csv", descriptive_header=TRUE, export=TRUE, skip_num=59) {
  dir <- readLines(pipe(paste("ls", directory)))
  pathname <- paste(directory, "/", sep="")
  FILENUM <- 0
  non_csv <- 0
  cv <- function(x) 100*((sqrt(var(x)/length(x)))/mean(x)) #define CV
  cat("processing . . .\n")
  for (step in 1:length(dir)) {
    filename <- paste(pathname, dir[step], sep="")
    shortname <- dir[step]; corename <- substring(shortname, 1, nchar(shortname)-3)
    csv_test <- substring(filename, nchar(filename)-2, nchar(filename))
    boo <- any(as.logical(grep(csv_test, "csv")))
    if (boo) {
      if (FILENUM == 0) FILENUM<-1 else FILENUM <- 2 # used below for define of data_array
      if (descriptive_header) { # option for running manually stripped files
        TABLE <- read.table(filename, skip=skip_num,sep=",".check.names=FALSE, header=TRUE, fill=TRUE, strip.white=TRUE)
        ROWS <- length(TABLE[[1]]) #GET TABLE LENGTH TO LEAVE OFF LAST ROW
        TABLE <- read.table(filename, skip=skip_num,sep=",".check.names=FALSE, header=TRUE, fill=TRUE, strip.white=TRUE, nrow=ROWS-1)}
      else {TABLE <- read.table(filename, skip=skip_num,sep=",".check.names=FALSE, header=TRUE, fill=TRUE, strip.white=TRUE) }
      cat(paste(filename, "\n"))
      TABLE$Ch1_Median <- log(TABLE$Ch1_Median,10) # log transform
      TABLE$Ch1_Median <- TABLE$Ch1_Median/as.numeric(summary((TABLE$Ch1_Median))[3]) #normalize
      TABLE$Ch1_SignalNoiseRatio <- log(TABLE$Ch1_SignalNoiseRatio,10)
      TABLE$Ch1_SignalNoiseRatio <- TABLE$Ch1_SignalNoiseRatio/as.numeric(summary((TABLE$Ch1_SignalNoiseRatio))[3])
      TABLE$Ch2_Median <- log(TABLE$Ch2_Median,10) # log transform
      TABLE$Ch2_Median <- TABLE$Ch2_Median/as.numeric(summary((TABLE$Ch2_Median))[3]) #normalize
      TABLE$Ch2_SignalNoiseRatio <- log(TABLE$Ch2_SignalNoiseRatio,10)
      TABLE$Ch2_SignalNoiseRatio <- TABLE$Ch2_SignalNoiseRatio/as.numeric(summary((TABLE$Ch2_SignalNoiseRatio))[3]) #normalize
      tablef <- factor(TABLE$Name) #create factor levels
      # IgE data (channel 1)
      E_sum_I <- tapply(TABLE$Ch1_Median, tablef, mean)
      E_sum_SNR <- tapply(TABLE$Ch1_SignalNoiseRatio, tablef, mean)
      E_sum_CV <- tapply(TABLE$Ch1_Median, tablef, cv)
      E_sum_N <- tapply(TABLE$Ch1_Median, tablef, length)
      # IgG4 data (channel 2)
      G_sum_I <- tapply(TABLE$Ch2_Median, tablef, mean)
      G_sum_SNR <- tapply(TABLE$Ch2_SignalNoiseRatio, tablef, mean)
      G_sum_CV <- tapply(TABLE$Ch2_Median, tablef, cv)
      G_sum_N <- tapply(TABLE$Ch2_Median, tablef, length)
      # ratio data
      EtoG_I <- E_sum_I/ G_sum_I
      EtoG_SNR <- E_sum_SNR/ G_sum_SNR
      #quartz(); plot(sum_I, sum_CV, main=paste(corename, "raw data"))
    }
  }
}
```

(Continued)

TABLE E1. (Continued)

```

data_sum <- data.frame(subject=corename,element=row.names(E_sum_I),IgE_int=E_sum_I, IgE_SNR=E_sum_SNR, E_CV=E_sum_CV,
E_N=E_sum_N, IgG_int=G_sum_I, IgG_SNR=G_sum_SNR, G_CV=G_sum_CV, G_N=G_sum_N, Ratio_I=EtoG_I, Ratio_SNR=EtoG_SNR)
cluster_sum <- data.frame(Ratio_SNR=EtoG_SNR)
if (FILENUM==1) {data_array <- data_sum}
else {data_array <- rbind(data_array, data_sum)}
else non_csv <- non_csv + 1
}
cat(paste("total number of files:", paste(step, "\n")))
cat(paste("total number of csv files:", paste(step - non_csv, "\n")))
#g <- ggobi(data_array)
if (export) {
  exportpath <- paste(readLines(pipe("pwd")), "/", sep = "")
  write.csv(data_array, paste(exportpath, exportname, sep = ""))
  cat(paste("writing file", paste(exportpath, paste(exportname, "\n"), sep = "")))
}
}

BATCH(paste(current_dir, "/flinterman_data/VISIT_1", sep = ""), exportname="flinterman_data/V1_DATA", skip_num=60)
BATCH(paste(current_dir, "/flinterman_data/Negative_controls", sep = ""), exportname="flinterman_data/CTRL_DATA", skip_num=60)
v1 <- read.csv("flinterman_data/V1_DATA")
norm <- read.csv("flinterman_data/CTRL_DATA")
v1.clean <- subset(v1, E_CV<5); v1.clean <- subset(v1.clean, E_N == 4) #low CV and peptide elements only
norm.clean <- subset(norm, E_CV<5); norm.clean <- subset(norm.clean, E_N == 4) #low CV and peptide elements only

#now run loops to generate 'near neighbor' normalization
v1.nearE <- array(); norm.nearE <- array(); v1.nearG <- array(); norm.nearG <- array()
for (x in 1:length(v1.clean$IgE_SNR)){v1.nearE[x]<-median(c(v1.clean$IgE_SNR[x-1],v1.clean$IgE_SNR[x],v1.clean$IgE_SNR[x+1]))}
v1.nearE[length(v1.nearE)] <- 1 # remove NA from last value
for (x in 1:length(v1.clean$IgG_SNR)){v1.nearG[x]<-median(c(v1.clean$IgG_SNR[x-1],v1.clean$IgG_SNR[x],v1.clean$IgG_SNR[x+1]))}
v1.nearG[length(v1.nearG)] <- 1 # remove NA from last value
for (x in 1:length(norm.clean$IgE_SNR)){norm.nearE[x]<-median(c(norm.clean$IgE_SNR[x-1],norm.clean$IgE_SNR[x],norm.clean$IgE_SNR[x+1]))}
norm.nearE[length(norm.nearE)] <- 1 # remove NA from last value
for (x in 1:length(norm.clean$IgG_SNR)){norm.nearG[x]<-median(c(norm.clean$IgG_SNR[x-1],norm.clean$IgG_SNR[x],norm.clean$IgG_SNR[x+1]))}
norm.nearG[length(norm.nearG)] <- 1 # remove NA from last value

#create data frames
v1.near.EnG <- data.frame(element = v1.clean$element, subject = v1.clean$subject, E_SNR = v1.nearE, G_SNR = v1.nearG)
norm.near.EnG <- data.frame(element = norm.clean$element, subject = norm.clean$subject, E_SNR = norm.nearE, G_SNR = norm.nearG)

#summary plot showing mean signal and cutoff value using log transformed SNR
quartz()
elementf <- factor(v1.clean$element)
v1.Esum.mean.419 <- tapply(v1.clean$IgE_SNR, elementf, mean)
plot(log(v1.Esum.mean.419, 2), type="n", ylim=c(0,6), ylab="IgE SNR",main="all peptides low CV data")
lines(log(v1.Esum.mean.419, 2),col=4)
elementf <- factor(norm.clean$element)
norm.Esum.mean.419 <- tapply(norm.clean$IgE_SNR, elementf, mean)
lines(log(norm.Esum.mean.419, 2), col=2)

#summary plot showing mean signal and cutoff value using log transformed 'nearby' SNR
quartz()
elementf <- factor(v1.near.EnG$element)
v1.Esum.mean.nearby <- tapply(v1.near.EnG$E_SNR, elementf, mean)
v1.Gsum.mean.nearby <- tapply(v1.near.EnG$G_SNR, elementf, mean)
plot(log(v1.Esum.mean.nearby, 2), type="n", ylim=c(-1,3), ylab="Log Signal Noise Ratio",main="nearby analysis low CV data")
lines(log(v1.Esum.mean.nearby, 2), col=4)
#display IgG4 below axis
lines(-1*log(v1.Gsum.mean.nearby, 2), col=3)
elementf <- factor(norm.near.EnG$element)
norm.Esum.mean.nearby <- tapply(norm.near.EnG$E_SNR, elementf, mean)
norm.Gsum.mean.nearby <- tapply(norm.near.EnG$G_SNR, elementf, mean)
lines(log(norm.Esum.mean.nearby, 2), col=2); lines(-1*log(norm.Gsum.mean.nearby, 2), col=2)
abline(v=c(205.5, 254.5), lty=3)

```

(Continued)

TABLE E1. (Continued)

```

text(102,2.8,"Ara h1", cex=2)
text(230, 2.8,"Ara h2", cex=2)
text(330, 2.8,"Ara h3", cex=2)
text(-9,-0.9,"IgG4"); text(-10,2.9,"IgE")

#set cutoff for IgE and IgG
Ecutoff <- quantile(log(norm.near.EnG$E_SNR,2), 0.995, type=8)
abline(h=as.numeric(Ecutoff), lty=3)
Gcutoff <- quantile(log(norm.near.EnG$G_SNR,2), 0.995, type=8)
abline(h=-1*as.numeric(Gcutoff), lty=3)
quartz(); hist(log(norm.near.EnG$E_SNR,2)); abline(v=as.numeric(Ecutoff), lty=3)
quartz(); hist(log(norm.near.EnG$G_SNR,2)); abline(v=as.numeric(Gcutoff), lty=3)
v1.E.logical <- array(); v1.G.logical <- array()
for (x in 1:length(v1.near.EnG$E_SNR)) {
  v1.E.logical[x] <- as.numeric(log(v1.near.EnG$E_SNR[x],2)) > as.numeric(Ecutoff)
  v1.G.logical[x] <- as.numeric(log(v1.near.EnG$G_SNR[x],2)) > as.numeric(Gcutoff) }
norm.E.logical <- array(); norm.G.logical <- array()
for (x in 1:length(norm.near.EnG$E_SNR)) {
  norm.E.logical[x] <- as.numeric(log(norm.near.EnG$E_SNR[x],2)) > as.numeric(Ecutoff)
  norm.G.logical[x] <- as.numeric(log(norm.near.EnG$G_SNR[x],2)) > as.numeric(Gcutoff) }

#create EvG plots of normals and V1
quartz()
double_pos <- sum(as.numeric(norm.G.logical & norm.E.logical))
E.pos <- sum(as.numeric(norm.E.logical)) - double_pos
G.pos <- sum(as.numeric(norm.G.logical)) - double_pos
plot(log(norm.near.EnG$E_SNR,2), log(norm.near.EnG$G_SNR,2), xlim=c(-1.6), ylim=c(-1.6), xlab="log IgE SNR", ylab="log IgG SNR",
  main="Normal controls n=6"); abline(v=as.numeric(Ecutoff),h=as.numeric(Gcutoff), lty=3)
text(5.5,-0.5, round(E.pos/length(norm.E.logical) * 100, 2), cex=1.5)
text(5.5,5.5, round(double_pos/length(norm.E.logical) * 100, 2), cex=1.5)
text(-0.5,5.5, round(G.pos/length(norm.E.logical) * 100, 2), cex=1.5)

quartz()
double_pos <- sum(as.numeric(v1.G.logical & v1.E.logical))
E.pos <- sum(as.numeric(v1.E.logical)) - double_pos
G.pos <- sum(as.numeric(v1.G.logical)) - double_pos
plot(log(v1.near.EnG$E_SNR,2), log(v1.near.EnG$G_SNR,2), xlim=c(-1.6), ylim=c(-1.6), xlab="log IgE SNR", ylab="log IgG SNR", main="Peanut
  sensitized n=24")
abline(v=as.numeric(Ecutoff),h=as.numeric(Gcutoff), lty=3)
text(5.5,-0.5, round(E.pos/length(v1.E.logical) * 100, 2), cex=1.5)
text(5.5,5.5, round(double_pos/length(v1.E.logical) * 100, 2), cex=1.5)
text(-0.5,5.5, round(G.pos/length(v1.E.logical) * 100, 2), cex=1.5)

#create logical data.frame and plot for cutoff sum
#visit 1

v1.logical <- data.frame(element = v1.near.EnG$element, subject = v1.near.EnG$subject, Epos = v1.E.logical, Gpos = v1.G.logical)
elementf <- factor(v1.logical$element)
v1.logical.Esum <- tapply(v1.logical$Epos, elementf, sum)
quartz(); plot(v1.logical.Esum/length(levels(v1.logical$subject)), type="h", main="Peanut sensitized IgE", ylab="% positive")
lines(v1.logical.Esum/length(levels(v1.logical$subject)), col=4)
abline(v=c(205.5, 254.5), lty=3)
text(102,2.8,"Ara h1", cex=2)
text(230, 2.8,"Ara h2", cex=2)
text(330, 2.8,"Ara h3", cex=2)
text(-9,-0.9,"IgG4"); text(-10,2.9,"IgE")

v1.logical.Gsum <- tapply(v1.logical$Gpos, elementf, sum)
quartz(); plot(v1.logical.Gsum/length(levels(v1.logical$subject)), type="h", main="Peanut sensitized IgG", ylab="% positive")
lines(v1.logical.Gsum/length(levels(v1.logical$subject)), col=3)
abline(v=c(205.5, 254.5), lty=3)
text(102,2.8,"Ara h1", cex=2)
text(230, 2.8,"Ara h2", cex=2)
text(330, 2.8,"Ara h3", cex=2)
text(-9,-0.9,"IgG4"); text(-10,2.9,"IgE")

```

(Continued)

TABLE E1. (Continued)

#normal controls

```
norm.logical <- data.frame(element = norm.near.EnG$element, subject = norm.near.EnG$subject, Epos = norm.E.logical, Gpos = norm.G.logical)
elementf <- factor(norm.logical$element)
norm.logical.Esum <- tapply(norm.logical$Epos, elementf, sum)
quartz(); plot(norm.logical.Esum/length(levels(norm.logical$subject)), type="h", main="Normals IgE", ylab="% positive")
lines(norm.logical.Esum/length(levels(norm.logical$subject)), col=4)
```

```
norm.logical.Gsum <- tapply(norm.logical$Gpos, elementf, sum)
quartz(); plot(norm.logical.Gsum/length(levels(norm.logical$subject)), type="h", main="Normals IgG", ylab="% positive")
lines(norm.logical.Gsum/length(levels(norm.logical$subject)), col=3)
```

#calculate percent of patients recognizing at least one epitope of Arah1, Arah2 or Arah3

```
element <- v1.logical$element
Arah1 <- element %in% grep("Ara_h_1", as.character(element), value=TRUE)
A1.logical <- subset(v1.logical, Arah1)
subjectf <- A1.logical$subject
A1.percent.patients <- tapply(A1.logical$Epos, subjectf, sum)
```

```
Arah2 <- element %in% grep("Ara_h_2", as.character(element), value=TRUE)
A2.logical <- subset(v1.logical, Arah2)
subjectf <- A2.logical$subject
A2.percent.patients <- tapply(A2.logical$Epos, subjectf, sum)
```

```
Arah3 <- element %in% grep("Ara_h_3", as.character(element), value=TRUE)
A3.logical <- subset(v1.logical, Arah3)
#A3.logical <- subset(A3.logical, element != "Ara_h_3.128")
#A3.logical <- subset(A3.logical, element != "Ara_h_3.129")
#A3.logical <- subset(A3.logical, element != "Ara_h_3.130")
#A3.logical <- subset(A3.logical, element != "Ara_h_3.131")
#A3.logical <- subset(A3.logical, element != "Ara_h_3.132")
subjectf <- A3.logical$subject
A3.percent.patients <- tapply(A3.logical$Epos, subjectf, sum)
```

#explore relationship between epitopes and reactivity and create bins by epitope num and by combo score

```
subjectf <- factor(v1.logical$subject)
E.logical.sum <- tapply(v1.logical$Epos, subjectf, sum)
G.logical.sum <- tapply(v1.logical$Gpos, subjectf, sum)
patient.data <- read.csv(paste(current_dir, "/flinterman_data/score_key.csv", sep=""))
v1.patient <- subset(patient.data, VISIT == "v1")
v1.patient.sort <- v1.patient[ do.call(order, v1.patient) ,] #sort by subject

E.epitope.num <- data.frame(subject=names(E.logical.sum), E.epitope=E.logical.sum)
G.epitope.num <- data.frame(subject=names(G.logical.sum), G.epitope=G.logical.sum)
v1.E.epitope.num.sort <- E.epitope.num[ do.call(order, E.epitope.num) ,]
v1.G.epitope.num.sort <- G.epitope.num[ do.call(order, G.epitope.num) ,]
v1.patient <- cbind(v1.patient.sort, E.epitope.num=v1.E.epitope.num.sort$E.epitope, G.epitope.num=v1.G.epitope.num.sort$G.epitope)
quartz(); plot(v1.patient$E.epitope.num, v1.patient$COMBO, main="Reaction severity v. epitope diversity", ylab="Reaction score", xlab="# positive epitopes")
```

```
q1.epitope <- subset(v1.patient, E.epitope.num <= as.numeric(quantile(v1.patient$E.epitope.num, prob = 0.25)))
epitopeQ <- data.frame(Q=1, epitopes=q1.epitope$E.epitope.num, combo=q1.epitope$COMBO)
q2.epitope <- subset(v1.patient, E.epitope.num <= as.numeric(quantile(v1.patient$E.epitope.num, prob = 0.5)) & E.epitope.num >=
  as.numeric(quantile(v1.patient$E.epitope.num, prob = 0.25)))
epitopeQ <- rbind(epitopeQ, data.frame(Q=2, epitopes=q2.epitope$E.epitope.num, combo=q2.epitope$COMBO))
q3.epitope <- subset(v1.patient, E.epitope.num <= as.numeric(quantile(v1.patient$E.epitope.num, prob = 0.75)) & E.epitope.num >=
  as.numeric(quantile(v1.patient$E.epitope.num, prob = 0.5)))
epitopeQ <- rbind(epitopeQ, data.frame(Q=3, epitopes=q3.epitope$E.epitope.num, combo=q3.epitope$COMBO))
q4.epitope <- subset(v1.patient, E.epitope.num >= as.numeric(quantile(v1.patient$E.epitope.num, prob = 0.75)))
epitopeQ <- rbind(epitopeQ, data.frame(Q=4, epitopes=q4.epitope$E.epitope.num, combo=q4.epitope$COMBO))
```

```
quartz(); boxplot(epitopeQ$combo ~ epitopeQ$Q, ylim=c(0,15), main="Reaction score by epitope quartile", ylab="Combo reaction score")
```

(Continued)

TABLE E1. (Continued)

```

points(epitopeQ$combo ~ epitopeQ$Q, col=4, pch=19, cex=2)
epitope_extremes <- subset(epitopeQ, Q == 1 | Q == 4)

wilcox.test(subset(epitope_extremes,Q==1)$combo, subset(epitope_extremes,Q==4)$combo, alternative = "I")$p.value
text(4,14,"*",cex=2)

#comparison using sIgE
q1.sIgE <- subset(v1.patient, sIgE <= as.numeric(quantile(v1.patient$sIgE, prob = 0.25)))
sIgE_Q <- data.frame(Q=1, sIgE=q1.sIgE$sIgE, combo=q1.sIgE$COMBO)
q2.sIgE <- subset(v1.patient, sIgE <= as.numeric(quantile(v1.patient$sIgE, prob = 0.5)) & sIgE >= as.numeric(quantile(v1.patient$sIgE,
  prob = 0.25)))
sIgE_Q <- rbind(sIgE_Q, data.frame(Q=2, sIgE=q2.sIgE$sIgE, combo=q2.sIgE$COMBO))
q3.sIgE <- subset(v1.patient, sIgE <= as.numeric(quantile(v1.patient$sIgE, prob = 0.75)) & sIgE >= as.numeric(quantile(v1.patient$sIgE, prob = 0.5)))
sIgE_Q <- rbind(sIgE_Q, data.frame(Q=3, sIgE=q3.sIgE$sIgE, combo=q3.sIgE$COMBO))
q4.sIgE <- subset(v1.patient, sIgE >= as.numeric(quantile(v1.patient$sIgE, prob = 0.75)))
sIgE_Q <- rbind(sIgE_Q, data.frame(Q=4, sIgE=q4.sIgE$sIgE, combo=q4.sIgE$COMBO))

quartz(); boxplot(sIgE_Q$combo ~ sIgE_Q$Q, ylim=c(0,15), main="Reaction score by sIgE quartile", ylab="Combo reaction score")
points(sIgE_Q$combo ~ sIgE_Q$Q, col=4, pch=19, cex=2)
sIgE_extremes <- subset(sIgE_Q, Q == 1 | Q == 4)

wilcox.test(subset(sIgE_extremes,Q==1)$combo, subset(sIgE_extremes,Q==4)$combo, alternative = "I")$p.value

#quartile by combo score
q1.combo <- subset(v1.patient, COMBO <= as.numeric(quantile(v1.patient$COMBO, prob = 0.25)))
q1.near.EnG <- subset(v1.near.EnG, subject == as.character(q1.combo$subject[1]))
q1.logical <- subset(v1.logical, subject == as.character(q1.combo$subject[1]))
for (x in 2:length(q1.combo$subject)) {
  q1.near.EnG <- rbind(q1.near.EnG, subset(v1.near.EnG, subject == as.character(q1.combo$subject[x])))
  q1.logical <- rbind(q1.logical, subset(v1.logical, subject == as.character(q1.combo$subject[x]))) }
quartz(); plot(log(q1.near.EnG$E_SNR,2), log(q1.near.EnG$G_SNR,2), main="Q1 reaction severity", xlab="log IgE SNR", ylab="log IgG SNR",
  xlim=c(-1,6), ylim=c(-1,6))
double_pos <- sum(as.numeric(q1.logical$Epos & q1.logical$Gpos))
E.pos <- sum(as.numeric(q1.logical$Epos)) - double_pos
G.pos <- sum(as.numeric(q1.logical$Gpos)) - double_pos
abline(v=as.numeric(Ecutoff),h=as.numeric(Gcutoff), lty=3)
text(5.5,-0.5, round(E.pos/length(q1.logical$Epos) * 100, 2), cex=1.5)
text(5.5,5.5, round(double_pos/length(q1.logical$Epos) * 100, 2), cex=1.5)
text(-0.5,5.5, round(G.pos/length(q1.logical$Epos) * 100, 2), cex=1.5)

q2.combo <- subset(v1.patient, COMBO <= as.numeric(quantile(v1.patient$COMBO, prob = 0.5)) & COMBO >=
  as.numeric(quantile(v1.patient$COMBO, prob = 0.25)))
q2.near.EnG <- subset(v1.near.EnG, subject == as.character(q2.combo$subject[1]))
q2.logical <- subset(v1.logical, subject == as.character(q2.combo$subject[1]))
for (x in 2:length(q2.combo$subject)) {
  q2.near.EnG <- rbind(q2.near.EnG, subset(v1.near.EnG, subject == as.character(q2.combo$subject[x])))
  q2.logical <- rbind(q2.logical, subset(v1.logical, subject == as.character(q2.combo$subject[x]))) }
quartz(); plot(log(q2.near.EnG$E_SNR,2), log(q2.near.EnG$G_SNR,2), main="Q2 reaction severity", xlab="log IgE SNR", ylab="log IgG SNR",
  xlim=c(-1,6), ylim=c(-1,6))
double_pos <- sum(as.numeric(q2.logical$Epos & q2.logical$Gpos))
E.pos <- sum(as.numeric(q2.logical$Epos)) - double_pos
G.pos <- sum(as.numeric(q2.logical$Gpos)) - double_pos
abline(v=as.numeric(Ecutoff),h=as.numeric(Gcutoff), lty=3)
text(5.5,-0.5, round(E.pos/length(q2.logical$Epos) * 100, 2), cex=1.5)
text(5.5,5.5, round(double_pos/length(q2.logical$Epos) * 100, 2), cex=1.5)
text(-0.5,5.5, round(G.pos/length(q2.logical$Epos) * 100, 2), cex=1.5)

q3.combo <- subset(v1.patient, COMBO <= as.numeric(quantile(v1.patient$COMBO, prob = 0.75)) & COMBO >=
  as.numeric(quantile(v1.patient$COMBO, prob = 0.5)))
q3.near.EnG <- subset(v1.near.EnG, subject == as.character(q3.combo$subject[1]))
q3.logical <- subset(v1.logical, subject == as.character(q3.combo$subject[1]))
for (x in 2:length(q3.combo$subject)) {

```

(Continued)

TABLE E1. (Continued)

```

q3.near.EnG <- rbind(q3.near.EnG, subset(v1.near.EnG, subject == as.character(q3.combo$subject[x])))
q3.logical <- rbind(q3.logical, subset(v1.logical, subject == as.character(q3.combo$subject[x]))) }
quartz(); plot(log(q3.near.EnG$E_SNR,2), log(q3.near.EnG$G_SNR,2), main="Q3 reaction severity", xlab="log IgE SNR", ylab="log IgG SNR",
  xlim=c(-1,6), ylim=c(-1,6))
double_pos <- sum(as.numeric(q3.logical$Epos & q3.logical$Gpos))
E.pos <- sum(as.numeric(q3.logical$Epos)) - double_pos
G.pos <- sum(as.numeric(q3.logical$Gpos)) - double_pos
abline(v=as.numeric(Ecutoff),h=as.numeric(Gcutoff), lty=3)
text(5.5,-0.5, round(E.pos/length(q3.logical$Epos) * 100, 2), cex=1.5)
text(5.5,5.5, round(double_pos/length(q3.logical$Epos) * 100, 2), cex=1.5)
text(-0.5,5.5, round(G.pos/length(q3.logical$Epos) * 100, 2), cex=1.5)

q4.combo <- subset(v1.patient, COMBO >= as.numeric(quantile(v1.patient$COMBO, prob = 0.75)))
q4.near.EnG <- subset(v1.near.EnG, subject == as.character(q4.combo$subject[1]))
q4.logical <- subset(v1.logical, subject == as.character(q3.combo$subject[1]))
for (x in 2:length(q4.combo$subject)) {
  q4.near.EnG <- rbind(q4.near.EnG, subset(v1.near.EnG, subject == as.character(q4.combo$subject[x])))
  q4.logical <- rbind(q4.logical, subset(v1.logical, subject == as.character(q4.combo$subject[x]))) }
quartz(); plot(log(q4.near.EnG$E_SNR,2), log(q4.near.EnG$G_SNR,2), main="Q4 reaction severity", xlab="log IgE SNR", ylab="log IgG
  SNR",xlim=c(-1,6), ylim=c(-1,6))
double_pos <- sum(as.numeric(q4.logical$Epos & q4.logical$Gpos))
E.pos <- sum(as.numeric(q4.logical$Epos)) - double_pos
G.pos <- sum(as.numeric(q4.logical$Gpos)) - double_pos
abline(v=as.numeric(Ecutoff),h=as.numeric(Gcutoff), lty=3)
text(5.5,-0.5, round(E.pos/length(q4.logical$Epos) * 100, 2), cex=1.5)
text(5.5,5.5, round(double_pos/length(q4.logical$Epos) * 100, 2), cex=1.5)
text(-0.5,5.5, round(G.pos/length(q4.logical$Epos) * 100, 2), cex=1.5)

# V4 data

BATCH(paste(current_dir, "flinterman_data/VISIT_4", sep=""), exportname="flinterman_data/V4_DATA", skip_num=60)
v4 <- read.csv("flinterman_data/V4_DATA")
v4.clean <- subset(v4, E_CV<5); v4.clean <- subset(v4.clean, E_N == 4) #low CV and peptide elements only
norm.clean <- subset(norm, E_CV<5); norm.clean <- subset(norm.clean, E_N == 4) #low CV and peptide elements only
#now run loops to generate 'near neighbor' normalization
v4.nearE <- array(); norm.nearE <- array(); v4.nearG <- array(); norm.nearG <- array()
for (x in 1:length(v4.clean$IgE_SNR)){v4.nearE[x]<-median(c(v4.clean$IgE_SNR[x-1],v4.clean$IgE_SNR[x],v4.clean$IgE_SNR[x+1]))}
v4.nearE[length(v4.nearE)] <- 1 # remove NA from last value
for (x in 1:length(v4.clean$IgG_SNR)){v4.nearG[x]<-median(c(v4.clean$IgG_SNR[x-1],v4.clean$IgG_SNR[x],v4.clean$IgG_SNR[x+1]))}

v4.nearG[length(v4.nearG)] <- 1 # remove NA from last value
#create data frame

v4.near.EnG <- data.frame(element = v4.clean$element, subject = v4.clean$subject, E_SNR = v4.nearE, G_SNR = v4.nearG)

#summary plot showing mean signal and cutoff value using log transformed 'nearby' SNR

quartz()
elementf <- factor(v4.near.EnG$element)
v4.Esum.mean.nearby <- tapply(v4.near.EnG$E_SNR, elementf, mean)
v4.Gsum.mean.nearby <- tapply(v4.near.EnG$G_SNR, elementf, mean)
plot(log(v4.Esum.mean.nearby, 2), type="n", ylim=c(-1,3), ylab="Log Signal Noise Ratio",main="nearby analysis low CV data")
lines(log(v4.Esum.mean.nearby, 2), col=4)
#display IgG4 below axis

lines(-1*log(v4.Gsum.mean.nearby, 2), col=3)
#create logical data.frame and plot for cutoff sum
#visit 4
v4.E.logical <- array(); v4.G.logical <- array()
for (x in 1:length(v4.near.EnG$E_SNR)) {
  v4.E.logical[x] <- as.numeric(log(v4.near.EnG$E_SNR[x],2)) > as.numeric(Ecutoff)
  v4.G.logical[x] <- as.numeric(log(v4.near.EnG$G_SNR[x],2)) > as.numeric(Gcutoff) }
v4.logical <- data.frame(element = v4.near.EnG$element, subject = v4.near.EnG$subject, Epos = v4.E.logical, Gpos = v4.G.logical)
elementf <- factor(v4.logical$element)

```

(Continued)

TABLE E1. (Continued)

```
v4.logical.Esum <- apply(v4.logical$Epos, elementf, sum)
quartz(); plot(v4.logical.Esum/length(levels(v4.logical$subject)), type="h", main="Peanut sensitized IgE", ylab="% positive")
lines(v4.logical.Esum/length(levels(v4.logical$subject)), col=4)
v4.logical.Gsum <- apply(v4.logical$Gpos, elementf, sum)
quartz(); plot(v4.logical.Gsum/length(levels(v4.logical$subject)), type="h", main="Peanut sensitized IgG", ylab="% positive")
lines(v4.logical.Gsum/length(levels(v4.logical$subject)), col=3)

#create matrices for contour plots
v1v4E <- array(0, dim=c(21,20))
v1v4G <- array(0, dim=c(21,20))
for (x in 1:length(v1.logical.Esum)) {
  v1v4E[as.numeric(v1.logical.Esum[x])+1, as.numeric(v4.logical.Esum[x])+1] <- v1v4E[as.numeric(v1.logical.Esum[x])+1,
as.numeric(v4.logical.Esum[x])+1] + 1
  v1v4G[as.numeric(v1.logical.Gsum[x])+1, as.numeric(v4.logical.Gsum[x])+1] <- v1v4G[as.numeric(v1.logical.Gsum[x])+1,
as.numeric(v4.logical.Gsum[x])+1] + 1
}
quartz(); contour(x=seq(0, 21,len=nrow(v1v4E)), y=seq(0, 20, len=ncol(v1v4E)), nlevels=as.numeric(max(v1v4E)), v1v4E, lwd=2, drawlabels = FALSE,
col=4, main="Stability of IgE epitopes over 20 months", ylab="visit 4 sum positive by epitope", xlab="visit 1 sum positive by epitope"); abline(0,1,
col=4)
quartz(); contour(x=seq(0, 21,len=nrow(v1v4G)), y=seq(0, 20, len=ncol(v1v4G)), nlevels=as.numeric(max(v1v4G)), v1v4G, lwd=2, drawlabels =
FALSE, col=3, main="Stability of IgG epitopes over 20 months", ylab="visit 4 sum positive by epitope", xlab="visit 1 sum positive by epitope");
abline(0,1, col=3)
chisq.test(v1.logical.Esum, v4.logical.Esum, simulate.p.value=TRUE); chisq.test(v1.logical.Esum, rnorm(419), simulate.p.value=TRUE)
chisq.test(v1.logical.Gsum, v4.logical.Gsum, simulate.p.value=TRUE); chisq.test(v1.logical.Esum, rnorm(419), simulate.p.value=TRUE)

# create heatmap image

# first need to create column data.frame
subjectf <- factor(v1.near.EnG$subject)
elementf <- factor(v1.near.EnG$element)
tempE <- by(v1.near.EnG$E_SNR, list(subjectf, elementf), sum)
tempG <- by(v1.near.EnG$G_SNR, list(subjectf, elementf), sum)
temp <- log(tempE/tempG, 2)
v1.cluster <- array(as.numeric(temp), dim=c(24, 419))
row.names(v1.cluster) <- dimnames(temp)[[1]]
names(v1.cluster) <- dimnames(temp)[[2]]
quartz(); heatmap(as.matrix(v1.cluster), Colv=NA, cexCol = 0.1 + .1/log10(419), col=topo.colors(20), labCol = names(v1.cluster), scale="none")
rm("temp", "tempE", "tempG")
#end script
```