

Distinction between peanut allergy and tolerance by characterization of B-cell receptor repertoires

Anna Ehlers¹, Constance den Hartog Jager¹, André Knulst¹, and Henderikus Otten¹

¹Universitair Medisch Centrum Utrecht

November 20, 2020

Abstract

Background Specific IgE against a peanut 2S albumin (Ara h2 or 6) is the best predictor of clinically relevant peanut sensitization. However, sIgE levels of peanut allergic and those of peanut sensitized but tolerant patients partly overlap, highlighting the need for improved diagnostics to prevent incorrect diagnosis and consequently unnecessary food restrictions. Thus, we sought to explore differences in V(D)J gene transcripts coding for peanut 2S albumin-specific monoclonal antibodies (mAbs) from allergic and sensitized but tolerant donors. Methods 2S albumin-binding B-cells were single-cell sorted from peripheral blood of peanut allergic (n=6) and tolerant (n=6) donors sensitized to Ara h2 and/or 6 ([?] 0.1 kU/l) and non-atopic controls (n=5). Corresponding heavy and light chain gene transcripts were heterologously expressed as mAbs and tested for specificity to native Ara h2 and 6. HCDR3 sequence motifs were identified by Levenshtein distances and hierarchically clustering. Results The frequency of 2S albumin-binding B-cells was increased in allergic (median: 0.01%) compared to tolerant (median: 0.006%) and non-atopic donors (median: 0.0015%, p=0.008). The majority of mAbs (74%, 29/39) bound specifically to Ara h2 and/or 6. Non-specific mAbs (9/10) were mainly derived from non-atopic controls. In allergic donors, 89% of heavy chain gene transcripts consisted of VH3-family genes, compared with only 54% in sensitized but tolerant and 63% of non-atopic donors. Additionally, certain HCDR3 sequence motifs were associated with allergy or tolerance upon hierarchical clustering of their Levenshtein distances. Conclusions HCDR3 sequence motifs associated with allergy or tolerance may support correct diagnosis of patients with suspected peanut allergy.

Distinction between peanut allergy and tolerance by characterization of B-cell receptor repertoires

Running title: Specific B-cell repertoires in peanut allergy

* Anna M. Ehlers, MSc^{1,2}

Constance F. den Hartog Jager, BSc^{1,2}

André C. Knulst^{1,2}, MD PhD

Henny G. Otten¹, PhD

¹Center for Translational Immunology, University Medical Center Utrecht, Utrecht University, Netherlands

²Department of Dermatology/Allergology, University Medical Center Utrecht, Utrecht University, Netherlands

***Corresponding author:**

Anna M. Ehlers, MSc

University Medical Center Utrecht

Laboratory of Translational Immunology
Heidelberglaan 100, 3584 CX, Utrecht, The Netherlands
Tel: +31 88 7559024, Fax: +31 88 7555404
Email: a.m.ehlers@umcutrecht.nl

Acknowledgments

We would like to thank all participants for their efforts and dedication in this study. Additionally, we would like to thank Anouska Michelsen-Huisman, Anne van Dijk, Ans Lebens and Mehran Alizadeh Aghdam for their support in the patient inclusion and Michael Drefahl, Rowena Melchers and Miriam Katsburg for their technical support. For critical discussion and reading of the manuscript, we would like to thank Melanie Abram, Edward Knol and Femke van Wijk. For financial support, we want to acknowledge EUROIMMUN AG, Lübeck, Germany.

Author contributions

AE, CHJ, HO: experimental design, AE, AC: patient recruitment, AE, CHJ: experimental performance, AE, AC, HO: data analysis and interpretation, AE: drafting manuscript, CHJ, AC, HO: critical revision of the manuscript, all authors approved the final manuscript

Funding

This research was financially supported by EUROIMMUN AG, Lübeck, Germany. They were not involved in the overall design of the study.

Conflict of Interest

The research position of A. Ehlers is partially funded by EUROIMMUN AG, Lübeck, Germany. Other authors have no conflicts of interest to declare.

Abstract

Background

Specific IgE against a peanut 2S albumin (Ara h2 or 6) is the best predictor of clinically relevant peanut sensitization. However, sIgE levels of peanut allergic and those of peanut sensitized but tolerant patients partly overlap, highlighting the need for improved diagnostics to prevent incorrect diagnosis and consequently unnecessary food restrictions. Thus, we sought to explore differences in V(D)J gene transcripts coding for peanut 2S albumin-specific monoclonal antibodies (mAbs) from allergic and sensitized but tolerant donors

Methods

2S albumin-binding B-cells were single-cell sorted from peripheral blood of peanut allergic (n=6) and tolerant (n=6) donors sensitized to *Ara h2* and/or 6 ([?] 0.1 kU/l) and non-atopic controls (n=5). *h 2 and/or 6* ([?] 0.1 kU/l). Corresponding h h eavy and light chain gene transcripts were heterologously expressed as mAbs and tested for specificity to native Ara h2 and 6. HCDR3 sequence motifs were identified by Levenshtein distances and hierarchically clustering.

Results

The frequency of 2S albumin-binding B-cells was increased in allergic (median: 0.01%) compared to tolerant (median: 0.006%) and non-atopic donors (median: 0.0015%, p=0.008). The majority of mAbs (74%, 29/39) bound specifically to Ara h2 and/or 6. Non-specific mAbs (9/10) were mainly derived from non-atopic controls. In allergic donors, 89% of heavy chain gene transcripts consisted of VH3-family genes, compared with only 54% in sensitized but tolerant and 63% of non-atopic donors. Additionally, certain HCDR3 sequence motifs were associated with allergy or tolerance upon hierarchical clustering of their Levenshtein distances.

Conclusions

HCDR3 sequence motifs associated with allergy or tolerance may support correct diagnosis of patients with suspected peanut allergy.

Keywords: clinically irrelevant sensitization, food allergy diagnostics, monoclonal antibodies, peanut allergy, VH family gene usage

Word count : 3500

Figure/Table count: 3 Figures and 3 Tables

Abbreviations

BAT Basophil activation test

DBPCFC double-blind placebo-controlled food challenge

HCDR3 complementary-determining region 3 of the heavy chain

mAbs monoclonal antibodies

OD optical density

PBMCs peripheral blood mononuclear cells

RBL rat basophil leukemia cells

sIgE specific IgE

V(D)J rearranged variable (V), diversity (D) and joining (J) gene segments

VH V (variable) gene of the heavy chain

Introduction

Current food allergy diagnostics comprise careful history, skin prick tests, measuring specific IgE (sIgE) and double-blind placebo-controlled food challenges (DBPCFC) as the gold standard. However, DBPCFCs are burdensome for the patient, costly and require dedicated hospital facilities^{1, 2}. On the other hand, current technologies to measure sIgE detect both clinically relevant and irrelevant sensitization, potentially leading to incorrect diagnosis and consequently unnecessary food restrictions^{3, 4}. Clinically relevant peanut sensitization is associated with sIgE against the major peanut allergens belonging to the 2S albumin family, Ara h2 and 6. In previous studies in adults, 100% positive predictive values for sIgE against Ara h2 and 6 were found using sIgE positivity thresholds of respectively 1.75 kU/l and 1.8 kU/l. Specific IgE levels below these thresholds, however, overlapped between allergic and tolerant subjects, preventing precise diagnosis^{5, 6}.

The occurrence of clinically irrelevant sensitization to Ara h2 and 6 might be explained by differences in peanut (Ara h2 and 6) specific antibody repertoires. These differences may include the presence of non-IgE antibodies blocking the binding of clinically relevant epitopes by sIgE of tolerant patients. Moreover, differences may also be based on antibody affinity and/or epitope recognition patterns. So far, no clear differences between allergic and tolerant subjects were observed by epitope mapping approaches⁷. These evaluations may have been hampered by the use of patient sera consisting of polyclonal IgE antibodies. Sera of allergic subjects usually contain mixtures of antibodies recognizing both clinically relevant and irrelevant epitopes. On the other hand, sera from tolerant subjects might contain antibodies recognizing clinically relevant epitopes with insufficient affinity for successful FcεRI receptor crosslinking, as well as antibodies recognizing clinically irrelevant epitopes⁸. Hence, deep analysis of monoclonal antibodies (mAbs) from specific B-cells may provide more insights into differences in specific antibody repertoires between allergic and tolerant subjects.

To this end, we analyzed gene sequences encoding the variable region of peanut 2S albumin-specific mAbs from 6 allergic and 6 sensitized but tolerant adults. In particular, sequence motifs of the HCDR3 region, the

most important region for recognizing antigens⁹, were hierarchically clustered. Clustering of HCDR3 regions resulted in four motifs exclusively present in allergic donors and three motifs associated with tolerance.

Methods

Blood samples (100 ml) were drawn from allergic (n=6) and tolerant (n=6) adults ([?]18 years) sensitized to Ara h2 and/or 6 ([?]0.1 kU/l¹⁰). Classification was achieved by DBPCFC or convincing history (n=2) of tolerated peanut ingestion. Non-atopic donors (n=5) served as reference group. The present study was approved by the ethical committee of the University Medical Center Utrecht (No. 17-945) and informed consent was given by all participants.

2S albumins isolated from roasted peanuts were used for tetramer formation with Streptavidin-APC or -PE^{11, 12}. B-cells double positive for tetramer-PE/APC staining were single cell sorted and the V(D)J gene transcript of the heavy and corresponding light chain were amplified as described previously^{11, 13, 14}. Quality checked sequences of successfully amplified gene transcripts were evaluated with the IgBlast web interface¹⁵. pFUSEss vectors (Invivogen), containing the IgE or IgG1 backbone, were used for antibody cloning and expression in HEK293 cells. Generated mAbs were tested for their specificity to nAra h2 and/or 6¹⁶ in comparison with their binding to transferrin. The ability of generated mAbs to induce degranulation was determined by CD63 upregulation of re-sensitized and stimulated human basophils.

Descriptive gene lineage analysis consisted of isotype distribution, mutational status and VH-family gene usage^{17, 18}. Sequence motifs were identified by calculating Levenshtein distances of HCDR3 regions combined with hierarchically clustering. HCDR3 sequences with five or less differences were defined as one motif. Detailed descriptions are shown in Supplementary File 1. These sequence data have been submitted to the GenBank database under submission number 2395667.

Results

Specific IgE levels overlapped between allergic and tolerant donors

In order to study differences in antibody repertoires between peanut allergic and peanut sensitized (IgE) but tolerant patients, blood was drawn from peanut allergic (n=6, age: 27-41) and tolerant (n=6, age: 27-63) donors sensitized to Ara h2 and/or 6 ([?]0.1 kU/l). Specific IgE levels to Ara h2 ranged from 1.0- 72 kU/l in the allergic and from 0-1.7 kU/l in the tolerant group. Specific IgE levels to Ara h 6 were lower compared with sIgE levels to Ara h2 and ranged from 0-13 kU/l in the allergic and from 0-0.9 kU/l in the tolerant group. Mono-sensitization to either Ara h2 or 6 was detected in both groups, although it occurred more often within the tolerant group (allergic: Pt-01, tolerant: Pt-02, 03, 07) as shown in Table 1.

Frequency of peanut 2S albumin-binding B-cells was significantly higher in allergic donors

As a first step, the frequencies of 2S albumin-binding B-cells were compared between peanut allergic patients, peanut sensitized but tolerant patients, and non-atopic controls. 2S albumin-binding B-cells, double-positive for allergen-tetramer staining, were defined as putatively specific, and their frequency was expressed as percentage of total B-cells acquired from the respective sample. While the frequency in allergic patients (median: 0.01%, 95% CI: 0.005-0.164) was only slightly increased compared to the tolerant patients (median: 0.006%, 95% CI: 0.0016-0.014), the frequency was significantly elevated compared with non-atopic controls (median: 0.002%, 95% CI: 0.0004-0.004, p=.008). Those 2S albumin-binding B-cells were single-cell sorted and served as source for heavy and corresponding light chain gene transcript amplification. No correlation was found between the frequency of 2S albumin-binding B-cells and the number of 2S albumin-binding B-cells from which the heavy and the corresponding light chain gene transcripts were successfully amplified and sequenced. A high number of V(D)J gene transcripts were successfully amplified from the B-cells of patients 4 and 6 – both allergic – while none or only one V(D)J gene transcript was successfully amplified from B-cells of patient 2 (tolerant) and 9 (allergic), respectively. For both of these patients, only a small number of 2S albumin-binding B-cells were sorted (Pt-2: 8, Pt-9: 12), which may exclude amplification failure as a reason for obtaining such a low number of successfully amplified gene transcripts. Frequencies, number of sorted 2S albumin-binding B-cells and successfully amplified corresponding V(D)J gene transcripts are

shown in Figure 111Sorted B-cells/amplified corresponding V(D)J gene transcripts - allergic: Pt-1 (32/49), Pt-4 (63/192), Pt-5 (8/96), Pt-6 (61/192), Pt-9 (1/12), Pt-10 (19/96); tolerant: Pt-2 (0/8), Pt-3 (7/108), Pt-7 (13/40), Pt-8 (7/36), Pt-11 (36/144), Pt-12 (17/72); non-atopic: NA-1 (21/50), NA-2 (14/32), NA-3 (7/36), NA-4 (8/48), NA-5 (12/24). Taken together, those successfully sequenced gene transcripts - heavy chain: 280, light chain: 221 - provide a good basis for further analyses.

2S albumin-binding B-cells show specific binding to Ara h2 and/or 6

In order to verify the specificity of identified 2S albumin-binding B-cells, 42 different sets of heavy and corresponding light chain gene transcripts were heterologously expressed as IgE or IgG1 mAbs. The majority of these mAbs (81%, 34/42), mostly derived from allergic and tolerant patients, showed specific binding to Ara h2 and 6 at 10 $\mu\text{g/ml}$ or lower in a concentration-dependent manner (Figure 2A). These heterologously expressed mAbs originated from 2S albumin-binding IgM+ (41%, 14/34), IgG1+ (3%, 1/34), IgG2+ (23%, 8/34), IgE+ (3%, 1/34), IgA1+ (18%, 6/34) and IgA2+ (9%, 3/34) B-cells. While the two mAbs with relatively high optical density (OD) values originated from IgG1+ and IgA1+ B-cells, mAbs with relatively low OD values originated predominately from IgM+ B-cells. About half of the specific mAbs (58%, 11/19) tested for separate binding to Ara h 2 and 62 or 6 recognized both allergens with comparable OD values. Eight mAbs exclusively recognized either Ara h2 or 6 (21%, 4/19 Ara h2 and 21%, 4/19 Ara h 6) (Figure 2B). The remaining antibodies without specific binding to Ara h2 or 6 at 10 $\mu\text{g/ml}$ originated predominately from IgM+ B-cells of non-atopic controls (88%, 7/8). These data confirm the specificity of most selected 2S albumin-binding B-cells, supporting the relevance of the gene analyses described in the following paragraphs.

Selected 2S albumin-specific mAbs can induce degranulation

As IgE B-cell memory is potentially stored in IgG1+ memory B-cells and to a lesser extent in B-cells belonging to another isotype, the functionality of mAbs originated from non-IgE 2S albumin-binding B-cells was examined¹⁹⁻²¹. To determine mAb functionality, stripped human basophils were loaded with mixtures of randomly selected mAbs (IgE) and subsequently stimulated with Ara h2 and 6 (Figure 2C and Supplementary, Figure S4). Stimulated basophils showed individual CD63 upregulation upon stimulation when loaded with mixtures containing the mAb Pt-6-2 (Mix-1, 2, 3 and 6). Such responses were already achieved at low concentrations (5 ng/ml Ara h2/6). Moreover, the overall CD63 upregulation induced by Mix 3 and 6 was even higher than the CD63 upregulation induced by the anti-IgE control at all allergen concentrations, pointing to an additive effect by using mAb Pt-6-2 (Mix-1) in combination with additional mAbs. The functionality of mAb Pt-6-2 and the observed additive effect confirm the selection of specific and functional B-cells.

Allergic and tolerant donors possess more class-switched 2S albumin-binding B-cells than non-atopic controls

All successfully sequenced heavy chain gene transcripts were initially analyzed for isotype distribution, as class-switching is mostly accompanied by somatic hypermutation maturation upon antigen challenge. Overall, IgM was the most dominant isotype across all groups. Notably, its prevalence was much lower in allergic (47%) and tolerant (53%) donors than in the non-atopic reference group (94%), indicating more antigen-challenged and matured specific B-cells originated from allergic and tolerant patients. Moreover, the prevalence of IgA2-expressing B-cells was increased in the tolerant group (20%), while a comparable distribution was observed in allergic and non-atopic donors (2%), as shown in Figure 3A. This finding suggests potential protection by specific IgA in tolerant patients, as IgA is generally able to prevent mucosal antigen crossing²². Regarding IgE, only a small number of V(D)J gene transcripts derived from IgE+ B-cells were successfully amplified (allergic: 1; tolerant: 2), possibly due to the extreme low abundance (0.002-1%) of IgE+ B-cells within the circulation²³. Even though slightly higher proportions are known for IgG3+ (~1%) and IgG4+ B-cells (~0.75%)^{24, 25}, no V(D)J gene transcripts derived from IgG3+ and IgG4+ B-cells were amplified. Overall, the individual isotype distribution varied for each patient, e.g. 2S albumin-binding B-cells of the allergic patient 4 were predominantly IgM+ B-cells, while those of patient 1 were predominantly IgG+ B-cells with a shift to class-switched B-cells in peanut allergic and tolerant patients. This indicates that the 2S albumin-binding B-cells from allergic and tolerant patients were more

often matured and antigen-challenged B-cells compared with the non-atopic control group.

IgM+ B-cells from allergic donors partly show strong maturation by somatic hypermutations

The introduction of non-silent (mutations on amino acid level) somatic hypermutations can increase target specificity and affinity during B-cell maturation. As shown in Figure 3B, the median number of non-silent mutations was increased in VDJ heavy chain gene transcripts of class-switched B-cells and was the highest for IgA1+ and IgA2+ B-cells (IgA1: 5-11 mutations, IgA2: 10-16 mutations). The number of non-silent mutations in VDJ gene transcripts of IgM expressing B-cells, on the other hand, ranged from 0 to more than 10 in all study groups. Of note, a small number of IgM+ B-cells, mostly originating from the allergic group, consisted of highly mutated heavy chain gene transcripts with mutation numbers comparable to the other isotypes, pointing to potential maturation of 2S albumin-binding IgM+ B-cells by antigen challenge.

VH3-family genes are predominately used in gene transcripts of allergic donors

The variable region of the heavy chain gene transcript consists of recombined V, D and J genes. The V gene accounts for most nucleotides of the variable region and contributes, besides D and J genes, to the CDR3 region. Hence, heavy chain V gene lineage can greatly affect B-cell receptor specificity and affinity. In line with B-cell repertoires described in literature²⁶, VH3-family genes were observed to dominate across all groups, with a shift to higher proportions (89%, mainly IGHV3-30, VH3-23 and VH3-72) in allergic donors ($X^2(2, n=183)=23.67, p<.0001$). This significant difference was still present upon neglecting redundant sequences from the same donor, excluding a bias by proliferation of a certain 2S albumin-binding B-cells. In turn, IGHV4 genes accounted for a larger part of incorporated VH-genes in tolerant (21%) donors than in allergic ones (7%) (Table 2). Overall, VH-family gene usage differs between allergic and tolerant patients sensitized to Ara h2/6, supporting the hypothesis of differences in allergen specific antibody repertoires between allergic and tolerant patients.

An IgE+ B-cell is largely clonally related to an IgG1+ B-cell in allergic patients

In total, three VDJ gene transcripts were successfully amplified from IgE+ B-cells, representing the smallest isotype group when disregarding IgG3 and IgG4. Where two VDJ gene transcripts of IgE+ B-cells derived from one tolerant donor were not closely related to another amplified VDJ gene transcript derived from the same donor, the VDJ gene transcript of one IgE+ B-cell originating from an allergic donor (Pt-10) was nearly identical to the gene transcript of an IgG1+ B-cell derived from an unrelated allergic donor (Pt-04). These two sequences differed only in one silent mutation within the FR4 region and one non-silent mutation within the FR2 region, pointing towards clonal relationship between IgE+ and IgG1+ B-cells and conserved clones between unrelated donors (Figure 3C). This finding supports the hypothesis that the IgE memory is stored in IgG1+ B-cells and to lesser extent in B-cells of another isotype^{20, 21}.

HCDR3 sequence motifs are related to allergy or tolerance

As HCDR3 regions have a great impact on antigen binding, their characteristics were more deeply analyzed. Detailed analysis of physiochemical properties, such as amino acid distribution, aliphatic index, Boman index and Kidera factors 1-10, are shown in the Supplementary, Figure S5 and 6¹⁷. The Kidera factor 5, expressing the preference to form double-banded structures, was significantly higher for the HCDR3 region of allergic donors compared with that of tolerant donors ($p = 0.0226$, Bonferroni correction: non-significant, Figure 3D). Nevertheless, the mean Boman index, describing the theoretical ability to bind proteins, was comparable between HCDR3 regions of allergic, tolerant, and non-atopic donors. For deeper analysis, HCDR3 sequence motifs associated with peanut allergy and/or tolerance, were evaluated by Levenshtein distances ([?]5 replacements, deletions or insertions) and hierarchical clustering. Levenshtein distances were favored over clonal relationship analysis because of individual genetic variation regarding haplotype diversity, single nucleotide polymorphisms, gene copy number and preferred gene lineage²⁷. For this analysis, all successfully amplified and sequenced heavy chain gene transcripts were used irrespectively of successfully sequenced corresponding light chain gene transcripts. Overall, four unique HCDR3 sequence motifs were associated with peanut allergy and three with peanut tolerance (Figure 3E, Table 3). The most dominant motif was

‘CARDSSALEIYNRFDPW’ (motif 1), which was derived from 36 different B-cells belonging to three different allergic donors (Pt-1 (n=32), Pt-5 (n=3), Pt-6 (n=1)) including 32 nearly identical B-cells of Pt-1 (monoclonal proliferation). This motif was formed by VH3-30, DH3-3 and JH5 genes. Besides motifs exclusively related to either allergy or tolerance, five motifs were shared between allergic, tolerant, and non-atopic donors. It must be mentioned that one motif (CARNVFDGYWLWYV) associated with tolerance was only found in Pt-11 and no motif was shared between all allergic or tolerant donors. Heterologously expressed mAbs corresponding to these motifs showed specific binding to Ara h2 or 6, supporting the relevance of the identified motifs and their potential role in diagnosing patients with suspected peanut allergy.

Discussion

Specific IgE levels to Ara h2 and 6 between 0.1 and 1.8 kU/l overlap between peanut allergic and tolerant subjects^{5, 6}, risking inaccurate diagnosis, and therefore indicating the need for new diagnostic strategies. In the present study, we observed a preferential usage of VH3-family genes in peanut 2S albumin-specific B-cells from peanut allergic patients. Additionally, we identified 2S albumin-binding B-cells carrying HCDR3 sequence motifs either related to peanut allergy or tolerance. This finding may lead towards new diagnostic strategies able to discriminate between allergy and tolerance in sensitized patients with suspected peanut allergy.

Despite a large number of amplified heavy chain VDJ gene transcripts (n=280), only three of them belonged to IgE class-switched B-cells. This low abundance may be potentially explained by dominant class-switching to IgE in the tissue¹⁹, low BCR surface expression of plasmablast-like IgE+ B-cells²⁸⁻³⁰ and the extreme low abundance of IgE+ B-cells within the circulation²³. Nevertheless, the relevance of the B-cells sorted within the present study is supported by the close relationship between one heavy chain gene transcript from an IgE+ B-cell and a gene transcript from an IgG1+ B-cell derived from an unrelated peanut allergic donor. Additionally, some generated mAbs were able to slightly inhibit patient-derived serum IgE-binding to Ara h2 and 6 (Supplementary, Figure S3). Correspondingly, there is evidence that human IgE+ B-cells are predominately plasmablasts or plasma cells generated by sequential class-switching from IgG1+ B-cells (γ switch region remnants) and that the humoral IgE memory is contained in IgG+ memory B-cells²¹. Further evidence for this theory was obtained by clonal relationship analyses, since IgE+ B-cells were dominantly clonally related to IgG1+ B-cells, but also, to a lesser extent, to B-cells of other isotypes^{19, 20, 28}.

By contrast to the low number of IgE+ B-cells, IgM+ B-cells represented the largest isotype group, as detected in respectively 47, 53 and 94% of allergic, tolerant and non-atopic donors. Additionally, IgM+ B-cells from non-atopic donors, all specific for Ara h2 and/or 6, shared HCDR3 motifs (motif 4, 8, 10) with IgM+, IgG2+ and IgA1+ B-cells from allergic donors. Clonal relationships between IgA/G/M+ and IgE+ B-cells have been described for B-cells derived from gut tissues by Hoh and colleagues^{19, 31}. Combining the findings from the present study with the finding of Hoh and colleagues leads to the suggestion that non-atopic donors can potentially possess non-IgE antibodies with required specificity or affinity to theoretically induce an allergic reaction.

Moreover, tolerant donors showed a higher proportion of IgA+ B-cells and tended to have higher specific IgA serum levels than allergic donors (Supplementary, Figure S3). Allergen-specific IgA serum levels have been shown to be increased in peanut allergic subjects undergoing oral or sublingual immunotherapy compared with their baseline levels. Moreover, salivary IgA levels have been closely associated with the degree of tolerance induction confirmed by DBPCFCs, pointing to a protective role of specific IgA against mucosal allergen absorption^{22, 32}. Such a protective role is supported by increased intestinal permeability in the absence of IgA in mice³³. Taken together, these findings suggest a potential protective effect of allergen-specific IgA in the mucosa of sensitized but tolerant patients.

Regarding gene lineage, VH3-family genes were significantly more often used ($p < .0001$) in heavy chain gene transcripts of peanut 2S albumin-specific B-cells from allergic donors (89%) than in those from tolerant (54%) and non-atopic donors (63%). Previous studies on VH-gene usage of heavy chain gene transcripts in healthy donors showed VH3-family gene usage in 40 to 65% of the B-cells, which was comparable to the usage

in the tolerant and non-atopic groups in our study (Table 2)^{34, 35}. These findings suggest a shift towards VH3-family gene usage in 2S albumin specific B-cells from allergic donors. Contrary to our findings, other datasets of Ara h2 specific B-cells did not observe a similar shift in VH-gene usage^{11, 14}. These datasets, however, were generated from patients undergoing peanut oral immunotherapy with a focus on different time points during immunotherapy. This fact hampers the comparison between the present and previous Ara h2-related datasets. Moreover, conflicting results regarding preferred VH-gene usage were described for different food and respiratory allergies, with a shift to VH3 usage for anti-alpha Gal antibodies and anti-grass pollen Phl p6 and 11 antibodies^{36, 37}, indicating that the preferred VH-family gene usage may be allergen dependent. Moreover, the preferred usage might depend on the status of allergy or tolerance and can potentially be used for diagnostic purposes upon validation in a larger patient cohort. A validation study will also provide information about the number of detected 2S albumin-binding B-cells is sufficient for a diagnostic workflow.

Besides differences in VH-family gene usage, certain HCDR3 sequence motifs were associated with either peanut allergy or tolerance and appear to have the ability to discriminate between those two groups. The main HCDR3 sequence motif 'CARDSSALEIYNRFDPW' was associated with peanut allergy and derived from recombined VH3-30, DH3-3 and JH5 genes. Comparably, VH3-30*18 was incorporated in the VDJ gene transcript of clonally related IgE+ B-cells specific for Ara h2 in the study of Croote and colleagues²⁸. Additionally, a highly similar HCDR3 region (CAREGYESSGFDYW) to motif 6 (CAREGHY SSQ FDYW), associated with tolerance, has been described for peanut allergic subjects undergoing oral immunotherapy. Oral immunotherapy may shape the antibody repertoire towards repertoires present in tolerant subjects¹⁴. These comparisons support the potential of HCDR3 motifs in diagnosing peanut allergy and tolerance.

In conclusion, the dominant usage of VH3-family genes and the identification of HCDR3 sequence motifs related to either peanut allergy or tolerance may potentially lead to the development of new diagnostic strategies for subjects with suspected peanut allergy and sensitization to Ara h2 and/or 6. Validation of these HCDR3 sequence motifs in a larger patient cohort may be achieved using next-generated sequencing approaches, potentially combined with the sorting strategy of 2S albumin-binding B-cells presented in this study. Next-generation sequencing allows high-throughput and can be more easily implemented in diagnostic routine^{38, 39}.

References

1. Muraro A, Werfel T, Hoffmann-Sommergruber K, Roberts G, Beyer K, Bindslev-Jensen C, et al. EAACI Food Allergy and Anaphylaxis Guidelines: diagnosis and management of food allergy. *Allergy* 2014; 69; doi: 10.1111/all.12429.
2. Yanagida N, Sato S, Asaumi T, Ogura K, Ebisawa M. Risk Factors for Severe Reactions during Double-Blind Placebo-Controlled Food Challenges. *Int Arch Allergy Immunol* 2017; 172:173-82; doi: 10.1159/000458724.
3. Portnoy J. Appropriate Allergy Testing and Interpretation. *Mo Med* 2011; 108:339-43.
4. Yu W, Freeland D, Nadeau K. Food allergy: immune mechanisms, diagnosis and immunotherapy. *Nat Rev Immunology* 2016; 16:751-65; doi:10.1038/nri.2016.111 .
5. Klemans R, Broekman H, Knol E, Bruijnzeel-Koomen C, Otten H, Pasmans S, et al. Ara h2 is the best predictor for peanut allergy in adults. *J Allergy Clin Immunol Pract* 2013; 1:632-8.e1; doi: 10.1016/j.jaip.2013.07.014.
6. Kansen H, van Erp F, Knulst A, Ehlers A, Lyons S, Knol E, et al. Peanut allergy is accurately predicted in one third of adults with suspected peanut allergy using specific IgE to Ara h2 and Ara h 6. EAACI. London, 2020.
7. Ehlers A, Blankestijn M, Knulst A, Klinge M, Otten H. Can alternative epitope mapping approaches increase the impact of B-cell epitopes in food allergy diagnostics? *Clin Exp Allergy* 2019; 49:17-26; doi:

10.1111/cea.13291.

8. Bucaite G, Kang-Pettinger T, Moreira J, Gould H, James L, Sutton B, et al. Interplay between Affinity and Valency in Effector Cell Degranulation: A Model System with Polcalcin Allergens and Human Patient-Derived IgE Antibodies. *J Immunol* 2019; 203:1693-700; doi: 10.4049/jimmunol.1900509.
9. Xu J, Davis M. Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity* 2000; 13:37-45; doi: 10.1016/s1074-7613(00)00006-6.
10. Martinet J, Couderc L, Renosi F, Bobée V, Marguet C, Boyer O. Diagnostic Value of Antigen-Specific Immunoglobulin E Immunoassays against Ara h2 and Ara h 8 Peanut Components in Child Food Allergy. *Int Arch Allergy Immunol* 2016; 169:216-22; doi: 10.1159/000446181.
11. Patil S, Ogunniyi A, Calatroni A, Tadigotla V, Ruiter B, Ma A, et al. Peanut oral immunotherapy transiently expands circulating Ara h2-specific B cells with a homologous repertoire in unrelated individuals. *J Allergy Clin Immunol* 2015; 136:125-34; doi: 10.1016/j.jaci.2015.03.026.
12. Franz B, May K, Dranoff G, Wucherpfennig K. Ex vivo characterization and isolation of rare memory B cells with antigen tetramers. *Blood* 2011; 118:348-57; doi: 10.1182/blood-2011-03-341917.
13. Tiller T, Meffre E, Yurasov S, Tsuiji M, Nussenzweig M, Wardemann H. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J Immunol Methods* 2008; 329:112-24; doi: 10.1016/j.jim.2007.09.017.
14. Hoh R, Joshi S, Liu Y, Wang C, Roskin K, Lee J, et al. Single B cell Deconvolution of Peanut-specific Antibody Responses in Allergic Patients. *J Allergy Clin Immunol* 2016; 137:157-67; doi: 10.1016/j.jaci.2015.05.029.
15. Ye J, Ma N, Madden T, Ostell J. IgBLAST: An Immunoglobulin Variable Domain Sequence Analysis Tool *Nuc Acid Res* 2013; 41:W34-40; doi: 10.1093/nar/gkt382.
16. Koppelman S, De Jong G, Laaper-Ertmann M, Peeters K, Knulst A, Hefle S, et al. Purification and immunoglobulin E-binding properties of peanut allergen Ara h 6: evidence for cross-reactivity with Ara h2. *Clin Exp Allergy* 2005; 35:490-7; doi: 10.1111/j.1365-2222.2005.02204.x.
17. Margreitter C, Lu H-C, Townsend C, Stewart A, Dunn-Walters D, Fraternali F. BRepertoire: a user-friendly web server for analysing antibody repertoire data. *Nuc Acid Res* 2018; 46:W264-70; doi: 10.1093/nar/gky276.
18. Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, McWilliam H, et al. Clustal W and Clustal X version 2.0 *Bioinformatics* 2007; 23:2947-8; doi: 10.1093/bioinformatics/btm404.
19. Hoh R, Joshi S, Lee J, Martin B, Varma S, Kwok S, et al. Origins and clonal convergence of gastrointestinal IgE+ B cells in human peanut allergy. *Sci Immunol* 2020; doi: 10.1126/sciimmunol.aay4209.
20. Looney T, Lee J, Roskin K, Hoh R, King J, Glanville J, et al. Human B-cell isotype switching origins of IgE. *J Allergy Clin Immunol* 2016; 137:579-86; doi: 10.1016/j.jaci.2015.07.014.
21. Saunders S, Ma E, Aranda C, Curotto de Lafaille M. Non-classical B Cell Memory of Allergic IgE Responses. *Front Immunol* 2019; 10:715; doi: 10.3389/fimmu.2019.00715.
22. LaHood N, Patil S, Shreffler W, Virkud Y, Ruiter B, Tang H, et al. Ara h2 Specific IgA B Cell Repertoire Matures During Peanut Oral Immunotherapy. *J Allergy Clin Immunol* 2020; 145.
23. Jimenez-Saiz R, Ellenbogen Y, Bruton K, Spill P, Sommer D, Lima H, et al. BCR analysis of single-sorted, putative IgE+ memory B cells in food allergy: an ephemeral existence? *J Allergy Clin Immunol* 2019; 144:336-9; doi: 10.1016/j.jaci.2019.04.001.

24. Karagiannis P, Villanova F, Josephs D, Correa I, Van Hemelrijck M, Hobbs C, et al. Elevated IgG4 in patient circulation is associated with the risk of disease progression in melanoma. *Oncoimmunology* 2015; 4; doi: 10.1080/2162402X.2015.1032492.
25. Marsh-Wakefield F, Ashhurst T, Trend S, McGuire H, Juillard P, Zinger A, et al. IgG3 + B cells are associated with the development of multiple sclerosis. *Clin Trans Immunology* 2020; 9:e01133; doi: 10.1002/cti2.1133.
26. Brezinschek H, Foster S, Brezinschek R, Dorner T, Domiati-Saad R, Lipsky P. Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(-)/IgM+ B cells. *J Clin Invest* 1997; 99:2488-501; doi: 10.1172/JCI119433.
27. Watson C, Glanville J, Marasco W. The Individual and Population Genetics of Antibody Immunity. *Trends Immunol* 2017; 38:459-70; doi: 10.1016/j.it.2017.04.003.
28. Croote D, Darmanis S, Nadeau K, Quake S. High-affinity allergen-specific human antibodies cloned from single IgE B cell transcriptomes. *Science* 2018; 362:1306-9; doi: 10.1126/science.aau2599.
29. Karnowski A, Acatz-Straussberger G, Klockenbusch C, Achatz G, Lamers M. Inefficient processing of mRNA for the membrane form of IgE is a genetic mechanism to limit recruitment of IgE-secreting cells. *Eur J Immunol* 2006; 36:1917-25; doi: 10.1002/eji.200535495.
30. Yang Z, Sullivan B, Allen C. Fluorescent in vivo detection reveals that IgE(+) B cells are restrained by an intrinsic cell fate predisposition. *Immunity* 2012; 36:857-72; doi: 10.1016/j.immuni.2012.02.009.
31. Iweala O, Burks A. IgE producers in the gut expand the gut's role in food allergy. *Nat Rev Gastroenterol Hepatol* 2020; 17:384-6; doi:10.1038/s41575-020-0309-5 .
32. Kulis M, Saba K, Kim E, Bird J, Kamilaris N, Vickery B, et al. Increased peanut-specific IgA levels in saliva correlate with food challenge outcomes after peanut sublingual immunotherapy. *J Allergy Clin Immunol* 2012; 129:1159-62; doi: 10.1016/j.jaci.2011.11.045.
33. Strait R, Mahler A, Hogan S, Khodoun M, Shibuya A, Finkelman F. Ingested allergens must be absorbed systemically to induce systemic anaphylaxis. *J Allergy Clin Immunol* 2012; 127:982-9.e1; doi: 10.1016/j.jaci.2011.01.034.
34. Graichy M, Galson J, Kovaltsuk A, von Niederhausern V, Pachlopnik Schmid J, Recher M, et al. Maturation of the Human B-Cell Receptor Repertoire with Age SSRN Electronic Journal 2020; doi: 10.2139/ssrn.3518535.
35. Goldstein L, Chen Y-J, Wu J, Chaudhuri S, Hsiao Y-C, Schneider K, et al. Massively Parallel Single-Cell B-cell Receptor Sequencing Enables Rapid Discovery of Diverse Antigen-Reactive Antibodies *Commun Biol* 2019; 2:304; doi: 10.1038/s42003-019-0551-y.
36. Wang L, Radic M, Galili U. Human anti-Gal Heavy Chain Genes. Preferential Use of VH3 and the Presence of Somatic Mutations. *J Immunol* 1995; 155:1276-85.
37. Persson H, Karbalaeei Sadegh M, Greiff L, Ohlin M. Delineating the specificity of an IgE-encoding transcriptome. *J Allergy Clin Immunol* 2007; 120:1186-92; doi: 10.1016/j.jaci.2007.06.041.
38. Hung S, Meissner B, Chavez E, Ben-Neriah S, Ennishi D, Jones M, et al. Assessment of Capture and Amplicon-Based Approaches for the Development of a Targeted Next-Generation Sequencing Pipeline to Personalize Lymphoma Management. *J Mol Diagn* 2018; 20:203-14; doi: 10.1016/j.jmoldx.2017.11.010.
39. Cummings C, Peters E, Lacroix L, Andre F, Lackner M. The role of next-generation sequencing in enabling personalized oncology therapy. *Clin Transl Sci* 2016; 9:283-92; doi: 10.1111/cts.12429.

Tables

Table 1: Patient characteristics

			Sex	Age	Last reaction	Seve
Allergic group	Alle					
Pt-05	Pt-05	Pt-05	f	34	3 years	seven
Pt-09	Pt-09	Pt-09	f	29	1 year	mod
Pt-01	Pt-01	Pt-01	f	39	1.5 years	mod
Pt-04	Pt-04	Pt-04	m	27	7 years	mod
Pt-10	Pt-10	Pt-10	f	27	3 years	seven
Pt-06	Pt-06	Pt-06	m	41	2 years	mod
Tolerant group	Tole					
Pt-02	Pt-02	Pt-02	m	45	N/A	N/A
Pt-07	Pt-07	Pt-07	f	29	N/A	N/A
Pt-08	Pt-08	Pt-08	m	36	N/A	N/A
Pt-11	Pt-11	Pt-11	m	27	N/A	N/A
Pt-12	Pt-12	Pt-12	f	31	N/A	N/A
Pt-03	Pt-03	Pt-03	m	63	N/A	N/A

Table 2: Comparison of the VH-gene family usage between the present study, healthy controls and Ara h2-related datasets (OIT trails); (transcripts from allergic, n = 120; tolerant, n = 63; non-atopic, n = 52)

VH family	Graichy+ 2020 ³⁴	Goldstein 2019 ³⁵	Patil 2015 ¹¹	Hoh 2016 ¹⁴	Allergic group	Tolerant group	Non-atopic group
VH1 [%]	5-18	15-18	23	18	2	21	3.8
VH2 [%]	2	4	-	-	-	-	-
VH3 [%]	40-65	46-50	62	60	89++	54	63
VH4 [%]	25-35	22-25	15	15	6.3	21	29
VH5 [%]	1-2	5-7	-	-	1.3	2.9	3.8
VH7 [%]	0.5-1	0.05-1	-	-	1.3	-	-

+ Age group between 25-40 years used for comparison, most closely related to the average age of the study population; data for different subsets were combined

++dominant usage VH3-family genes in all individuals

Table 3: HCDR3 sequence motifs associated with peanut allergy or tolerance

bold: differences of HCDR3 sequences; italic: tolerant donors; bold patient numbers: indicate heterologous expression mAb, n. e. = not expressed

Motif	HCDR3 sequence	Antibodies	Isotype	Specificity
Allergy	Allergy	Allergy		
1	CARDSSALEIYNRFDPW	1 - Pt-1-32+ , Pt-5-3, Pt-5-4, Pt-5-8+ , Pt-6-43+	IgG2	Ara h2 + 6
2	CVKDRQQYSSRWLD	Pt-5-2 , Pt-5-5, Pt-6-32+ , Pt-6-42, Pt-6-45, Pt-6-47, Pt-6-48	IgA1	Ara h2 + 6

Motif	HCDR3 sequence	Antibodies	Isotype	Specificity
3	CASMDILAANTHFGM	PD-6-17+ , Pt-6-37, Pt-6-52+ , Pt-10-19+	IgG2	Ara h2 + 6
4	CARGLVGANFYYYM	PD-4-10+ , Pt-4-41, Pt-4-61	IgM	Ara h2
	CARGRSGATYYYYY	NA-4-5+	IgM	Ara h2 + 6
	CARGRAGPSYYYYY	PD-4-11+	IgM	
Tolerance	Tolerance	Tolerance		
5	CARNVFDGYWLVYW	<i>Pt-11-3, Pt-11-4,</i> Pt-11-14+ , <i>Pt-11-15,</i> <i>Pt-11-23,</i> <i>Pt-11-24,</i> <i>Pt-11-25,</i> <i>Pt-11-26</i>	IgA2	Ara h2 + 6
6	CAREGHYSSQFDYW	<i>Pt-8-3, Pt-8-4,</i> <i>Pt-8-5, NA-2-3</i>	IgM	n. e.
7	CARDYGGYPHAAF	DI-11	IgM	n. e.
	CTRDTGTYPHAAF	NA-3-1, Pt-3-5	IgM, IgG2	Ara h2 + 6
Shared	Shared	Shared		
8	CTRPYRAFNWAI	NA-1-15	IgM	n. e.
	CTRPYRAFNWAT	NA-1-17, NA-5-2+ (only Ara h2), NA-5-5+ , NA-5-11, Pt-6-19	IgM, IgA1	Ara h2 + 6
9	CARVSSSWHTEYW	<i>Pt-3-7, Pt-6-33,</i> Pt-6-46+ , Pt-6-53, Pt-6-60, Pt-6-61	IgA1	Ara h2 + 6
10	CARGIIDKYGMDVW	<i>Pt-6-38, Pt-7-6,</i> Pt-10-13, Pt-10-14	IgG2	n. e.
	CARE YYYGMDVW	NA-4-1	IgM	n. e.
	CART LGYGMDVW	<i>Pt-11-16</i>	IgM	n. e.
11	CTRGAVSYTRHFQ	FW-4-20+	IgM	Ara h2
	CVRGALAYTRHFQ	FW-4-34+	IgM	
	CVRGAMSNTRHFQ	FW-4-28+	IgM	Ara h2/6++
	CARGAMSYTRHFQ	FW-8, Pt-4-40, Pt-4-63+, <i>Pt-8-6</i>	IgM	Ara h2/6++
12	CAKAYGSGSYLFDYW	<i>FW-7-4,</i> NA-1-10+	IgM	Ara h2 + 6
	CARGGSGSYTFDY	NA-1-4, NA-1-6, NA-1-13	IgM	
	CARSGGSGSYTFDY	FW-10-6+	IgM	Ara h2 + 6

- + specificity to Ara h2/6 proven
- ++only specificity against mixture of Ara h2 and 6

Figure Legends

Figure 1 Frequency of 2S albumin-binding B-cells

Frequency of 2S albumin-binding B-cells, double-positive for allergen-tetramer (PE and APC) staining, expressed as percentage of the total B-cell number acquired from the respective sample, indicated with the median. Each frequency is labelled with the respective patient number and the number of successfully amplified corresponding heavy and light V(D)J gene transcripts in relation to sorted B-cells. Statistical significance was proven with the Kruskal-Wallis test and corrected for multiple testing using the Dunn's multiple comparison test (adjusted p-value); all = allergic, tol = tolerant, n-a = non-atopic

Figure 2 Specificity and functionality of heterologously expressed mAbs

A Concentration-dependent (1 to 10 $\mu\text{g/ml}$) binding of heterologously expressed mAbs ($n = 42$) to Ara h2 and 6 defined as OD values corrected for their binding to transferrin; filled dots indicate mAbs expressed as IgE and open dots mAbs expressed as IgG1; bold dash line is the cut-off for IgE mAbs and non-bold dash line the cut-off for IgG1 antibodies

B Binding of heterologously expressed mAbs to Ara h2 and 6 ($n = 24$) at 10 g/ml defined as OD values corrected for their binding to transferrin; filled dots indicate mAbs expressed as IgE and open dots mAbs expressed as IgG1, bold dash line is the cut-off for IgE mAbs and non-bold dash line the cut-off for IgG1 mAbs; black = allergic, grey = tolerant, light grey = non-atopic

C Upregulation of CD63 upon loading of human basophils with mAb mixture and subsequent stimulation with Ara h2 and 6 in a concentration-dependent manner (5 to 1000 ng/ml); the dots indicate the mean of duplicate measurements and their range; Mix 1 = Pt-6-2, Mix 2 = Pt-6-2 + Pt-6-17, Mix 3 = Pt-6-2 + Pt-3-2 + Pt-6-11, Mix 4 = Pt-6-17, Pt-3-2, Pt-6-11, Mix 5 = Pt-1-32, Pt-6-11, Pt-6-6, Pt-6-8, Mix 6 = NA-1-10 + Pt-4-28 + Pt-4-34 + Pt-6-2, Mix 7 = Pt NA-1-10 + Pt-4-28 + Pt-4-34, Mix 8 = Pt-1-32 + Pt-6-15 + Pt-6-21, stripped basophils before loading, Rituximab expressed as IgE; Supplementary, Figure S4: comparison to stimulation with cow's milk extract, dash line: threshold level set to 5%

Figure 3: Descriptive gene analysis and motif analysis of HCDR3 regions

Gene lineage analysis of successfully sequenced heavy chain gene transcripts defined as productive (allergic $n = 151$; tolerant $n = 68$; non-atopic, $n = 51$), irrespectively of successfully sequenced corresponding light chain gene transcript

A Isotype distribution split for allergic (all), tolerant (tol) and non-atopic (n-a) donors

B Numbers of non-silent mutations expressed with the median (95% CI) and split for allergic (all), tolerant (tol) and non-atopic (n-a) donors

C Clonotypes were defined as identical V gene, J gene and HCDR3 length and related clones were aligned using ClustalW2. The IgE mAb of Pt-10 was highly related to an IgG1 antibody of Pt-4.

D Tendency of non-redundant HCDR3 regions (allergic $n = 88$; tolerant $n = 55$; non-atopic, $n = 48$) to form double-banded structures on amino acid level expressed as Kidera factor 5 in allergic (all), tolerant (tol) and non-atopic (n-a) donors (median and 95% CI). Correction for multiple measurements (Bonferroni) resulted in no statistically significant difference. More physicochemical properties of the HCDR3 region are shown in the Supplementary, Figure S5 and 6.

E HCDR3 sequence motifs were analyzed by Levenshtein distances and hierarchical clustering (detailed description in the method section) and sequences with distances [?] 5 were grouped into one motif. More detailed description of the motifs is presented in Table 3.



