HHS Public Access

Author manuscript

Sci Transl Med. Author manuscript; available in PMC 2024 April 11.

Published in final edited form as:

Sci Transl Med. 2024 February 07; 16(733): eadi0673. doi:10.1126/scitranslmed.adi0673.

CD23^{+}IgG1^{+} memory B cells are poised to switch to pathogenic IgE production in food allergy

Miyo Ota[1,2,†,‡,*], **Kenneth B. Hoehn**[3,†,§], **Weslley Fernandes-Braga**[1,2,†], **Takayuki Ota**[4], **Carlos J. Aranda**[1,2,‖], **Sara Friedman**[1,2], **Mariana G.C. Miranda-Waldetario**[1,2], **Jamie Redes**[1,2,5], **Maria Suprun**[1,¶], **Galina Grishina**[1], **Hugh A. Sampson**[1], **Alefiyah Malbari**[6,#], **Steven H. Kleinstein**[3,7,8], **Scott H. Sicherer**[1], **Maria A. Curotto de Lafaille**[1,2,*]

[1]Jaffe Food Allergy Institute, Division of Allergy and Immunology, Department of Pediatrics, Icahn School of Medicine at Mount Sinai (ISMMS); New York, NY 10029, USA

[2]Precision Immunology Institute (PrIISM), and Department of Immunology and Immunotherapy, ISMMS; New York, NY. 10029, USA

[3]Department of Pathology, Yale School of Medicine; New Haven, CT 06520, USA

[4]Department of Dermatology, Janssen Research & Development LLC; San Diego, CA 92121, USA

[5]Graduate School of Biomedical Sciences, ISMMS; New York, NY 10029, USA

[6]Kravis Children's Hospital, Department of Pediatrics, ISMMS; New York, NY 10029, USA

[7]Department of Immunobiology, Yale School of Medicine; New Haven, CT 06520, USA

[*]Corresponding authors. miyo.ota@absinstitute.org or maria.lafaille@mssm.edu.
[†]These authors contributed equally to this work
[‡]Present address: Applied Biomedical Science Institute; San Diego, CA 92127, USA
[§]Present address: Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth; Hanover, NH 03755, USA
[‖]Present address: Allergy Research Group, Instituto de Investigación, Biomédica de Málaga; Málaga 29590, Spain
[¶]Present address: Janssen Research & Development; Spring House, PA 19002, USA
[#]Present address: Department of Pediatrics and Department of Medical Education, The University of Texas Austin Dell Medical School; Austin, TX, USA

Author contributions:

MO and MACL conceptualized the study and MO designed the experimental approach. MO performed flow cytometry analysis, cell cultures, memory cell isolations for scRNA-seq, and sorting and BCR sequencing of antigen-binding cells. WFB performed gene expression analysis of antigen-binding cells, cell cultures of purified memory populations, and analysis of mAbs. SF characterized mAb binding to peanut antigens. MMW and JR helped with cell cultures and flow cytometry. KBH performed the computational analysis of scRNA-seq (transcriptomics and BCR) with supervision from SHK. TO helped with BCR repertoire analysis. MS and GG performed epitope binding analysis with HAS supervision. SHS and AM provided overall clinical oversight for allergic (SHS) and non-allergic (AM) participant recruitment. CJA and TO contributed to critical discussions. MO, KBH, WFB and MACL wrote the manuscript. MACL supervised the project. All authors read and approved the manuscript.

Competing interests: SHK receives consulting fees from Northrop Grumman and Peraton. KBH receives consulting fees from Prellis Biologics. MACL received consulting fees from Genentech. MS is an inventor of technology for detection of egg allergies using epitope-specific antibodies, and receives license payments from AllerGenis, LLC though ISMMS. MS and TO are employees at Janssen Pharmaceuticals. HAS receives consulting fees from N-Fold, LLC, DBV Technologies, AbbVie, and Siolta Therapeutics; holds Stock options in N-Fold and DBV Technologies; is an unpaid member of the Board of Directors of AllerGenis; receives grants from the NIH/NIAID and AllerGenis; and receives royalties from Elsevier. SHS reports royalty payments from UpToDate and from Johns Hopkins University Press; grants to his institution from the NIH/NIAID, Food Allergy Research and Education, and Pfizer; and personal fees from the American Academy of Allergy, Asthma and Immunology as Deputy Editor of the Journal of Allergy and Clinical Immunology: In Practice.

Data and materials availability: All data associated with this study are in the paper or supplementary materials. Scripts used to perform analysis are available at https://doi.org/10.5281/zenodo.10069762. Single-cell RNA sequencing data of this manuscript is contained in BioProject PRJNA847159 and GEO accession GSE208235.

[8]Program in Computational Biology & Bioinformatics, Yale University; New Haven, CT 06511, USA

Abstract

Food allergy is caused by allergen-specific IgE antibodies, but little is known about the B cell memory of persistent IgE responses. Here we describe, in human pediatric peanut allergy, a population of $CD23^+IgG1^+$ memory B cells arising in type 2 immune responses that contains high affinity peanut-specific clones and generate IgE-producing cells upon activation. The frequency of $CD23^+IgG1^+$ memory B cells correlated with circulating concentrations of IgE in children with peanut allergy. A corresponding population of "type 2-marked" $IgG1^+$ memory B cells was identified in single cell RNA sequencing experiments. These cells differentially expressed interleukin (IL)-4 and IL-13 regulated genes, such as *FCER2*/*CD23*, *IL4R*, and germline *IGHE*, and carried highly mutated B cell receptors (BCRs). In children with high concentrations of serum peanut-specific IgE, high affinity B cells that bind the main peanut allergen Ara h 2 mapped to the population of "type 2-marked" $IgG1^+$ memory B cells and included clones with convergent BCRs across different individuals. Our findings indicate that $CD23^+IgG1^+$ memory B cells transcribing germline *IGHE* are a unique memory population containing precursors of high affinity pathogenic IgE-producing cells that is likely to be involved in the long-term persistence of peanut allergy.

One Sentence Summary:

A population of IgG^+ memory B cells in peanut-allergic individuals contains high affinity allergen-specific clones and is poised to switch to IgE.

INTRODUCTION

High affinity IgE antibodies are essential mediators of food allergy, a main cause of life-threatening anaphylaxis. Most food allergies develop in childhood and disproportionately affect children (1). A long-standing question in the allergy field is why allergies to some foods spontaneously resolve, while others persist. A key to understanding the evolution of food allergy may reside in the mechanisms that maintain the B cell memory underlying high affinity IgE responses. Murine studies demonstrated a lack of functional IgE^+ memory B cells and that high affinity pathogenic IgE-producing plasma cells develop from the sequential switching of affinity matured $IgG1^+$ memory B cells (2, 3).

Human IgE^+ memory B cells and plasma cells identified by flow cytometry were found to be increased in the blood of allergic individuals (4, 5). These IgE^+ memory B cells were atypical as they expressed low amounts of the B cell receptor (BCR) signaling molecule CD79b and of membrane immunoglobulin (4), which may lead to compromised BCR signaling and function. Using single-cell RNA sequencing (scRNA-seq) analysis, Croote and collaborators demonstrated that most circulating cells are plasma cells and the rare IgE^+ memory B cells that were observed had reduced membrane IgE expression (6). Thus, in human as in mice, the plasma cell stage is the predominant type of IgE^+ cell, and although there are scarce human IgE^+ memory B cells, their functionality is questionable. Several human studies are consistent with the generation of IgE^+ cells from non-IgE-producing

memory precursors (7, 8). Sequencing of the switch region of IgE^+ B cells identified footprints of the switch regions of genes encoding IgGs (4, 9). Further, prior work has identified shared BCR sequence specificities between IgE^+ and IgG^+ B cells (10, 11). Longitudinal analysis of allergen-specific IgG and IgE antibodies also demonstrated that allergen-specific IgG precedes and accompanies allergen-specific IgE production (12, 13).

Peanut is the food allergen most often associated with anaphylaxis, and peanut allergy is generally persistent compared to other food allergies such as egg and cow's milk allergy (14-17). Circulating concentrations of peanut-specific IgE are predictive of thresholds of reactivity to peanut ingestion and whether peanut allergy will persist, though the relationship is imperfect, and other factors may play a role. In general, however, high peanut-specific IgE concentrations (> 100 kU_A/L) are associated with reactivity to a small amount of the food and long-term allergy persistence and reactivity to a small amount of the food, whereas low peanut-specific IgE (< 5 kU_A/L) is associated with reactivity to a relatively higher amount of the food and to outgrowing peanut allergy (17-22). We hypothesized that the existence of high affinity peanut-specific IgG^+ memory B cells and their ability to undergo class switching to IgE are critical for allergy persistence in pediatric peanut allergy. Here we described, in pediatric peanut-allergic patients, a population of IgG^+ memory B cells that is marked by expression of interleukin (IL)-4 and IL-13 regulated genes, contains high affinity allergen-specific clones, and is poised to switch to IgE.

RESULTS

$CD23^+IgG^+$ memory B cells frequency correlates with circulating IgE in pediatric peanut-allergic individuals.

The expression of CD23, the low affinity receptor for IgE, was previously found to be increased in B lymphocytes of children with atopic dermatitis (AD) and asthma (23, 24), and the frequency of $CD23^+$ B cells was found to correlate with IgE concentrations (25). To characterize B cells in food allergy, we first analyzed the expression of CD23 in B lymphocytes from peripheral blood mononuclear cells (PBMCs) of 58 peanut-allergic and 13 non-allergic children. CD27 was used as a marker of memory B cells that are somatically mutated and likely to contain allergen-specific clones (26). Consistent with previous observations, we found a significantly ($p<0.001$) higher percentage of $CD23^+$ total, naïve, and memory B cells in peanut-allergic than in non-peanut-allergic children of a similar age range (Fig. 1A and fig. S1A for gating strategy). Possible correlations between $CD23^+$ B cell populations and plasma IgE antibody concentration were analyzed in a cohort of peanut-allergic children with a broad range of peanut-specific IgE (Fig. 1B and C, fig. S1B). $CD27^+$ memory B cell populations were defined by expression of IgM/IgD (IgM/IgD^+ unswitched, and IgM/IgD^-), and by IgG expression among IgM/IgD^- memory cells. A significant correlation was found between total IgE and $CD23^+IgM/IgD^+$ memory B cells ($p=0.004$), $CD23^+$ IgM/IgD^- memory B cells ($p<0.001$), and $CD23^+IgG^+$ memory B cells ($p<0.001$) (Fig. 1B). The concentrations of total IgE in circulation also correlated significantly ($p<0.001$) with peanut-specific IgE and with IgE specific for the main peanut allergen, Ara h 2 (fig. S2A and B). No correlation was found between total IgE

concentrations and the frequency of $CD23^+$ total B cells (fig. S2C) or the $CD23^-$ memory populations (Fig. 1C).

A significant ($p<0.001$) difference in total IgE concentrations was observed between children with a very high concentration of peanut-specific IgE (PsIgE) in the blood (PsIgE > 100 kU_A/L referred to as PsIgE>100, n=18) and those with peanut-specific IgE lower than 5 kU_A/L (referred to as PsIgE<5, n=16), whereas total IgE concentrations varied among peanut-allergic children with intermediate peanut-specific IgE (5<PsIgE<100) (fig. S1B). To identify further differences in CD23 expression in the two ends of the spectrum of IgE sensitization, we compared the frequency of $CD23^+$ B cells between peanut-allergic children with PsIgE>100, peanut-allergic children with PsIgE<5, and non-allergic children (n=13) (table S1). The three groups had similar age ranges and median age. Both peanut-allergic groups had a higher frequency of $CD23^+$ naïve, $CD23^+$ total memory, and $CD23^+$ IgM/IgD^+ memory B cells than the non-allergic groups. However, only the PsIgE>100 peanut-allergic group had a significantly higher frequency of $CD23^+$ IgM/IgD^- memory B cells ($p<0.001$) and $CD23^+$ IgG^+ memory B cells ($p=0.002$) (Fig. 1D and fig. S2D) without a significant ($p>0.05$) increase in total IgG^+ memory B cells (fig. S3A and B). Peanut-allergic children with low concentrations of peanut-specific IgE tended to have lower frequency of $CD27^+$ memory B cells than children with high concentrations of peanut-specific IgE (fig. S3A and B). And although the frequency of $CD23^+$ naïve B cells was similarly increased in peanut-allergic children with very high or very low concentrations of peanut-specific IgE, only highly sensitized children had a significantly higher frequency of $CD23^+CD27^+IgG^+$ memory B cells as percentage of total B cells or of $CD27^+$ memory B cells compared to non-allergic children ($p=0.0002$ in Fig. 1D and $p=0.004$ in fig. S3C).

A cluster of somatically mutated switched memory cells express *FCER2*, *IL4R,* and *IGHE.*

To further characterize B cell memory subsets, we performed 10X Genomics scRNA-seq of whole transcriptomes and BCRs on flow-sorted $CD27^+$ memory B cells from 5 peanut-allergic children with PsIgE>100 and 3 non-allergic children (table S2). We identified 10 clusters of memory B cells (numbered 0 to 9) and identified differentially expressed genes (DEG) in each cluster (Fig. 2A and B, fig. S4 and 5, data file S1). Although genes for antibody isotypes except *IGHE* were not used in clustering, BCR analysis identified clusters enriched in unswitched IgM^+ memory B cells (clusters 3 and 9), switched memory cells (clusters 1, 4, 5, and 8), or containing a mixture of isotypes (clusters 0, 6, and 7) (Fig. 2C, fig. S6). Clusters 0 and 2 share main DEGs such as *YPEL*, *NR4A2*, *H3F3B,* and *FOSB*. *CD83* expression marked clusters 0, 2, 6, and 9, and other main DEGs of cluster 2 included *NR4A1*, *FOS,* and *JUNB*. Cluster 3 DEGs included *FGR* and *CD1C* and share some markers of atypical memory B cells (27) (fig. S7). Cluster 8 DEGs included major histocompatibility (MHC) genes *B2M*, *HLA-B/C/E*, *HLA-DRB5*, and *CD74*. Other DEGs of clusters 8 and 7 included genes encoding ribosomal proteins and RNA-binding proteins. The small cluster 9, which was highly enriched for *IGHM* BCRs, was characterized by high expression of *EGR1/2/3* and *MYC*, suggesting recent activation. Using mass cytometry, Glass *et al.* characterized memory populations by expression of CD27, CD45RB, CD73, CD95, or CD11c (28). Our transcriptional analysis was done on $CD27^+$ memory B cells, and it did not distinguish *CD45* isoforms. We detected *CD73/NT5E* expression in only a small

percentage of cells in clusters 4, 5, and 6, and transcripts for *FAS/CD95* and *ITGAX/CD11c* were undetectable in most cells (fig. S7).

Given the close association between $CD23^+IgG^+$ memory B cells and peanut allergy (Fig. 1), we searched for a cluster that contained $CD23^+IgG^+$ memory B cells. Cluster 5 was marked by high expression of *FCER2* (the gene that encodes CD23) (Fig. 2B, fig. S7) and contained mostly IgG^+ and IgA^+ cells (Fig. 2C). *FCER2* is upregulated by IL-4 and IL-13 signaling in B cells, and importantly, clusters 5 DEGs included other IL-4 and IL-13 regulated genes such *IL4R*, *CD86*, and *IGHE*. Thus cluster 5 contains type 2-marked memory cells like the ones we recently described (29). Top DEGs of cluster 5 included *HOPX* (Fig. 2B), encoding a homeobox-domain protein previously found to be associated with IgG^+ memory B cells and naïve B cells (30), *S100A10*, encoding a calcium-binding Annexin A2 partner protein that regulates the activity of many transmembrane proteins (31), *ANXA2* (encoding Annexin A2), and genes involved in MHC II antigen processing/presentation, such as *HLA-DPB1, HLA-DPA1, HLA-DQA1, HLA-DMB,* and *SCIMP* (Fig. 2B). Cluster 5, 4 and 6 shared DEGs were *PDE4D*, encoding phosphodiesterase 4D, *TXNIP*, encoding thioredoxin interacting protein, *SELL*, encoding selectin B, and *ARID5B*, encoding a transcription regulator of B lymphocytes. DEGs shared between clusters 5 and 4 were *COCH*, encoding cochlin, and *TOX*, a transcription factor expressed by germinal center cells and $CD80^+CD73^+IgG1^+$ mouse memory cells (3). DEGs in the type 2-marked memory cluster recently identified (29) included *FCER2*, *IL4R*, *IL13RA1*, *RNGTT*, *HOPX*, *CD74*, *PARM1* and *SCIMP*. Cluster 5, and to a lesser extent clusters 4 and 6, expressed these genes, indicating that cluster 5 is a type 2-marked memory cluster in our data (fig. S7).

To further understand the biological processes shaping the memory B cells in cluster 5, we performed gene ontology enrichment analysis. The most-enriched gene ontologies for cluster 5 were related to the cytokine-mediated signaling pathway (GO: 009221), and antigen receptor-mediated signaling pathway (GO: 0050851) (table S3). We next determined whether memory B cells in cluster 5 from peanut-allergic individuals had different expression signatures than cluster 5 cells from non-allergic individuals (data file S2). We found that cells from peanut-allergic individuals in cluster 5 were enriched for cytokine-mediated signaling pathway signatures (GO: 009221) (table S4). Overlap with this gene ontology included genes involved in signal transduction through janus kinase (JAK)-signal transducer and activator of transcription (STAT) and JAK-insulin receptor substrate 2 (IRS2) (*JAK1, PIM1, JUNB, CDKN1B, GRB2, HNRNPF, PTPN6,* and *SOCS1*). Thus, cluster 5 in peanut-allergic individuals was enriched for transcriptional signatures consistent with cytokine signaling, MHC II receptor activity, and JAK signaling.

We further characterized cluster 5 using BCR repertoire sequencing. Cluster 5 was highly enriched for *IGHG*-associated BCRs (61.7% of cells), predominantly *IGHG1* (40.1% of cells), as well as the highest proportion of *IGHG4*-associated BCRs among all clusters (7.84%) (Fig. 2C). BCRs in cluster 5 were highly mutated, with average somatic hypermutation (SHM) frequency of 8.08%, the highest of any cluster (Fig. 2D). This was not due to the low frequency of IGHM cells: *IGHG1*, *IGHG2*, and *IGHG3* cells in cluster 5 were more mutated than their respective cell types in other clusters (fig. S8). Although true IgE cells were rare, we identified two B cells with *IGHE* BCRs, each from a different peanut-

allergic individual and both belonging to cluster 5 (fig. S8). Further, both had detectable *FCER2* expression and high frequencies of SHM (7.3% and 11.1%).

In addition to the immunoglobulin constant region associated with a BCR, B cells can produce non-coding transcripts from the constant immunoglobulin genes. These could be germline transcripts indicating priming for near-future class switching between isotypes, or a past rearrangement in the non-productive *IGH* allele (32-34). In the scRNA-seq experiment, we identified 178 B cells with normalized *IGHE* transcript values at least as high as the two B cells found with *IGHE* BCRs (Fig. 3A). These *IGHE*-transcribing B cells had non-*IGHE* constant regions associated with their BCRs, predominantly *IGHG1* or *IGHG4* (62.4% and 18.5% of *IGHE*-transcribing cells, respectively) and were primarily found in cluster 5 (57.3% of *IGHE*-transcribing cells, Fig. 3B), suggesting that cluster 5 contains IgG^+ memory B cells primed to switch to IgE.

Peanut-specific B cells are found in IgG^+ memory B cell populations.

We subsequently asked if $CD23^+IgG^+$ memory B cells contained peanut-specific clones, and whether their frequencies were higher in PBMCs from highly sensitized peanut-allergic children (PsIgE>100) than in those with low peanut-specific sensitization (PsIgE<5). Since *IL4R* expression was transcriptionally upregulated in cluster 5 (Fig. 2B), which contained *FCER2/CD23* expressing IgG^+ memory B cells, we defined a sorting gate that comprised all cells with detectable expression of CD23 or IL-4 receptor (IL-4R) and referred to this population as $CD23^+/IL\text{-}4R^+$. $CD23^+/IL\text{-}4R^+IgG^+$ memory B cells and $CD23^-IL\text{-}4R^-$ double negative IgG^+ ($CD23^-IL\text{-}4R^-IgG^+$) memory B cells were sorted and cultured at 1,000 cells per well over fibroblasts expressing CD40 ligand (CD40L) and B-cell activating factor (BAFF) with the addition of the cytokines IL-2, IL-4, IL-10, and IL-21 (Fig. 4A and fig. S9A). These conditions were optimized for B cell proliferation and plasma cell differentiation, which also allowed class switching to IgE. After 7 days, peanut-specific antibodies were measured in the culture supernatants, and cells were analyzed by flow cytometry.

Most wells of cells from non-allergic individuals were negative for peanut-specific IgG antibodies (Fig. 4B, left). In cultures from peanut-allergic children with low peanut-specific IgE (PsIgE<5), the frequency of positive wells for peanut-specific IgG were similar between $CD23^-IL\text{-}4R^-$ and $CD23^+/IL\text{-}4R^+IgG^+$ cell cultures (Fig. 4B, middle). Importantly, in cultures of cells of peanut-allergic children with high peanut-specific IgE (PsIgE>100), wells containing peanut-specific IgG antibodies were predominantly found among $CD23^+/IL\text{-}4R^+IgG^+$ memory cell cultures (Fig. 4B, right). In addition, IgG^+ memory B cells from peanut-allergic children with high peanut-specific IgE (PsIgE>100) proliferated more than cells from the other two groups (fig. S9B). Class switching to IgE was relatively low in these culture conditions; nevertheless, we detected the highest frequency of IgE^+ B cells in the cultures of $CD23^+/IL\text{-}4R^+IgG^+$ memory B cells from peanut-allergic children with high peanut-specific IgE (PsIgE>100) (fig. S9C). To further identify the IgG^+ memory B cells containing peanut-specific clones in highly sensitized peanut-allergic children (PsIgE>100), we sorted out $CD23^+/IL\text{-}4R^+IgG^+$ memory B cells into $CD23^+$ and $CD23^-IL\text{-}4R^+$ fractions (fig. S9D) and cultured them as described in Fig. 4A over fibroblasts expressing CD40L,

BAFF, and IL-21. In day 9 culture supernatants, peanut-specific IgG was detected in 93% of wells from $CD23^+IgG^+$ memory B cells but only in 12% of wells from $CD23^-IL\text{-}4R^+IgG^+$ memory B cells (Fig. 4C). Importantly, IgG antibodies to Ara h 2 were only found in supernatants of $CD23^+IgG^+$ memory B cells (Fig. 4C). Total IgE was significantly ($p<0.001$) higher in supernatants from $CD23^+IgG^+$ memory B cells than in supernatants from $CD23^-IL\text{-}4R^+$ (or $CD23^-IL\text{-}4R^-$) IgG^+ memory B cells (fig. S9E). Although optical density (OD) values for peanut-specific and Ara h 2-specific IgE were very low, they were highest in cultures of $CD23^+IgG^+$ memory B cells (fig. S9F and G). To summarize, in highly sensitized peanut-allergic children, peanut-specific IgG memory clones are enriched in a $CD23^+$ population with the greatest ability to switch to IgE among IgG^+ memory B cells.

Sorted Ara h 2-specific $IgG1^+$ B cells are highly mutated and transcribe *FCER2* and germline *IGHE*.

To determine at the single-cell level if peanut-specific memory B cells were contained within an IgG^+ population with the characteristics of the cells described in cluster 5, we isolated B cells that bound the main peanut allergen, Ara h 2, from peanut-allergic patients (35-37). Ara h 2-binding B cells were single-cell sorted from PBMCs of 8 peanut-allergic individuals with PsIgE>100 and 5 peanut-allergic individuals with PsIgE<5 using fluorescent Ara h 2 multimers. For comparison, we sorted diphtheria toxin (DT)-binding B cells from peanut-allergic and non-allergic children, as these cells are expected to be found in most children due to widespread vaccination (table S5, fig. S10A). The frequency of Ara h 2 binding B cells was significantly ($p=0.02$) higher in peanut-allergic individuals with PsIgE>100 than in peanut-allergic individuals with PsIgE<5, whereas Ara h 2 binders were undetectable in non-allergic samples. DT binders were detected at similar frequencies among the three groups (fig. S10B).

To characterize the BCRs of Ara h 2-binding B cells, heavy and light chain BCR genes were sequenced from cDNA synthesized from sorted single cells. In total, we obtained heavy chain sequences from 69 Ara h 2-binding B cells and 174 DT-binding B cells, 86.4% of which had at least one associated light chain sequence. Ara h 2 binders from children with PsIgE>100 were predominantly IgG, and Ara h 2 binders from children with PsIgE<5 showed similar numbers of IgM and IgG whereas DT binders were predominantly IgG in individuals with PsIgE<5 and non-allergic individuals (fig. S10B). Heavy chain sequences of Ara h 2-binding IgG^+ B cells from peanut-allergic individuals were significantly ($p<0.001$) more mutated than DT-binding IgG^+ B cells from both peanut-allergic and non-allergic individuals (Fig. 5A). By contrast, most Ara h 2- and DT-binding IgM^+ cells were unmutated (Fig. 5A). Furthermore, Ara h 2-binding IgG^+ B cells from peanut-allergic individuals were significantly more likely to be of the IgG1 sub-isotype (85.7% of *IGHG* cells, Fig. 5B), compared to DT-binding $IgG1^+$ B cells from peanut-allergic (55.1% of $IGHG^+$ cells, $p = 0.0015$ by chi-square test) or from non-allergic individuals (52% of $IGHG^+$ cells, $p = 0.006$ by chi-square test).

To determine if Ara h 2-binding $IgG1^+$ cells belonged to the $CD23^+IgG^+$ memory population characterized through scRNA-seq (cluster 5; Fig. 2), we used real-time polymerase chain reaction (PCR) to quantify *FCER2/CD23*, germline *IGHE* (*GLT-IGHE*)

and total-*IGHE* transcripts in all antigen-binding IgG1$^+$ cells from peanut-allergic individuals with PsIgE>100 and from non-allergic individuals (Fig. 5C). *FCER2/CD23* expression was detected in most Ara h 2-binding IgG1$^+$ cells from peanut-allergic individuals, and its average expression was significantly higher in Ara h 2-binding IgG1$^+$ cells than DT-binding IgG1$^+$ cells from peanut-allergic (p=0.0028) or non-allergic individuals (p=0.0150). Importantly, total *IGHE* and germline *IGHE* transcripts were detected in most Ara h 2-binding cells but few DT-binding cells. These results show that Ara h 2 binding IgG$^+$ B cells from peanut-allergic individuals have defining characteristics of type 2-marked IgG$^+$ cells from cluster 5: high somatic hypermutation, predominant IgG1 constant regions, and transcription of *FCER2/CD23* as well as *IGHE*. The fact that Ara h 2 binding IgG1$^+$ memory cells transcribe germline *IGHE* is particularly relevant, as germline transcription is an essential initial step for class switch recombination (33).

High affinity convergent Ara h 2 specific BCR sequences were identified across different individuals.

BCRs within individual humans are highly diverse, and identification of similar sequences among different individuals may be indicative of convergent antigen-driven selection. To identify convergent BCRs specific for Ara h 2, we performed clustering based on IGH amino acid sequence similarity using the BCR sequencing data from Ara h 2- and DT-binders, and from all memory B cells profiled in the 10X scRNA-seq experiment. We refer to B cells with highly similar BCRs in multiple individuals as “convergent sequence families.”

Among cells isolated by Ara h 2 binding, monoclonal antibodies derived from a convergent sequence family of 2 IgM$^+$ memory cells (P5AC4, P5EC1) from two peanut-allergic individuals, as well as an antibody derived from an unrelated IgM$^+$ Ara h 2 binder (P2CC6), displayed very low binding to Ara h 2 by enzyme-linked immunosorbent assay (ELISA). Of these, only P2CC6 had some very low reactivity against Ara h 1 and Ara h 3 (fig. S11A). The heavy chains of these three IgM$^+$ cells had low SHM frequencies (between 0 and 2.1%).

Among Ara h 2-binding IgG$^+$ cells, we identified a group of convergent sequences composed of three IgG1$^+$ B cells (P4EC10, P4AC5, and P6EC2) using ambiguously IGHV3-30*18 or IGHV3-30-5*01 paired with IGHJ6*02, as well as IGKV3-20*01 and either IGKJ2*01 or IGKJ2*03 (data file S3). These cells were from two different peanut-allergic individuals and share between 73.4% and 78.2% amino acid identity with a previously described (PA13P1E10) anti-Ara h 2 IgE plasma cell (6) (Fig. 6A, PA13P1E10 family). Further, all three heavy chains were highly mutated, with SHM frequencies between 9.7% and 14.9%. To confirm whether the cells from this convergent group of sequences bound peanut antigens, we generated recombinant human monoclonal IgG1 antibodies. In ELISA and biolayer interferometry assays, we found that these antibodies bound with high affinity to Ara h 2 (K_D (M): <1.0E-12 for P4EC10) and cross-reacted with Ara h 1 and Ara h 3 (Fig. 6B, fig. S11B, fig. S12A and B).

A second convergent family of Ara h 2-binding IgG sequences mapped back to cluster 5 in the scRNA-seq data (Fig. 2 and 3). It contained three cells from two peanut-allergic individuals: one cell from cluster 5 of the 10X scRNA-seq (4C5G1) and two sorted Ara h

2 binder cells (P1BC6 and P1DC5, Fig. 6A, 4C5G1 family). Importantly, cell 4C5G1 from scRNA-seq cluster 5 had detectable transcripts of both *FCER2* and *IL4R*. All three cells in this cluster were used with IGHV1-8*01 and IGHJ4*02, as well as IGKV4-1*01 and either IGKJ3*02 or IGKJ2*01 (data file S3). These cells shared 73.0% to 88.6% amino acid identity in the heavy chain. Furthermore, all three heavy chain sequences had high SHM frequencies (between 11.5% and 14.2%) and all three bound Ara h 2 with high affinity (K_D (M): 3.42E-11 for 4C5G1; 1.13E-9 for P1BC6; 9.09E-11 for P1DC5) and cross-reacted with peanut antigens Ara h 1 and Ara h 3 (Fig. 6C, fig. S12A and B, fig. S13A). We identified three additional sequence families containing at least two Ara h 2-binding cells each, though these families did not span across multiple individuals. From each of these families, a single monoclonal IgG1 antibody was generated: P1BC5, P1EC5, and P6AC7 (data file S3). These three antibodies bound Ara h 2 with high affinity (fig. S12A).

The reactivity of the Ara h 2-specific monoclonal antibodies described above was also tested using a sequential linear epitope assay (38, 39). We found that all identified members of the 4C5G1 convergent family bound to the same Ara h 2 epitope (Ara h 2.008, fig. S13B). The 4C5G1 antibody also bound less intensely to a different epitope, which was also bound by members of the PA13P1E10 family (Ara h 2.019, fig. S13B). Members of the PA13P1E10 family bound mostly to three sequential epitopes (Ara h 2.018/019/020) and less intensely to Ara h 2.008 (fig. S13C). Outside from the convergent families, P1BC5 bound to Ara h 2.019, P1EC5 did not bind to any linear epitope, and P6AC7 bound to Ara h 2.008 and less to Ara h 2.019 (fig. S13B). Importantly, Ara h 2.008 and Ara h 2.019 are the two main linear epitope regions recognized by IgE antibodies in peanut-allergic patients (38) (fig. S13C). To summarize, we found that B cell clones that bind the main allergen Ara h 2 with high affinity were predominantly somatically-mutated, type 2-marked IgG1 memory cells that transcribed germline *IGHE*, contained convergent BCR sequences, and shared epitope recognition with IgE antibodies.

DISCUSSION

Using single-cell transcriptomics and functional analysis, we identified a population of somatically-mutated IgG^+ memory B cells characterized by expression of IL-4 and IL-13 regulated genes, *CD23/FCER2*, *IL4R*, and *IGHE*, that was composed mostly of $IgG1^+$ and $IgG4^+$ cells. This IgG^+ memory B cell population is very similar to the type 2-marked IgG^+ memory B cells that we recently described and showed were increased in adults with atopic disease (29), and to the MBC2 population described in the accompanying manuscript by Koenig *et al.* (40). Two other recent manuscripts described B cell populations transcriptionally related to the type 2-marked IgG^+ memory cells, but their BCR identity and link to IgE memory were not investigated (41, 42).

The frequency of $CD23^+IgG^+$ memory B cells in PBMC correlated with circulating IgE concentrations in a peanut-allergic pediatric population, which supports a relationship between these memory B cells and IgE. Expression of *IL4R*, in at least part of the $CD23^+$ population, suggests a readiness to respond to IL-4, and transcription of *IGHE* is particularly striking as germline *IGHE* transcription is essential to initiate class switch recombination to IgE (33). CD23 expression specifically marks the IgG^+ memory B cell population

containing peanut-specific clones in highly sensitized peanut-allergic individuals, as IgG^+ memory B cells that expressed IL-4R but not CD23 did not contain peanut-specific clones. IgG memory cells from highly sensitized peanut-allergic children produced the most IgE cells among IgG memory populations.

In agreement with the findings from the in vitro assays, IgG^+ memory B cells that bound the main peanut allergen Ara h 2 with high affinity were predominantly $IgG1^+$ cells expressing *FCER2* and germline *IGHE*, thus belonging to the *FCER2*$^+$ *IGHE*$^+$ IgG^+ population identified in cluster 5 of the scRNA-seq analysis. In line with our model of IgE cell differentiation developed from mouse studies (2, 3, 7, 43), we proposed that allergen-specific IgG^+ memory B cells expressing *CD23*/*FCER2* and germline *IGHE* are precursors of pathogenic IgE^+ plasma cells in highly sensitized peanut-allergic individuals.

Pathway analysis of differentially expressed genes of the *IGHE*$^+$ memory cluster (cluster 5) demonstrated enhanced expression of genes related to cytokine-mediated signaling and antigen receptor-mediated signaling. Within cluster 5, the cytokine-mediated signaling pathway was increased in peanut-allergic samples compared to non-allergic samples. Interestingly, the expression of *JAK1*, *GRB2*, *PTPN6,* and *SOCS1* was higher in peanut-allergic samples. This is consistent with differential activation of IL-4R-mediated JAK1-STAT6 signaling in cells from peanut-allergic individuals, which promotes class switching to IgE (33), as well as JAK1-IRS2 signaling, which induces survival and proliferation of B cells through growth factor receptor bound protein 2 (Grb2)-mediated activation of phosphoinositide 3-kinase (PI3K) and extracellular signal-regulated kinase 1/2 (ERK1/2) (44, 45). The increased expression of inhibitors *PTPN6* and *SOCS1* also supports the higher activity of the IL-4R pathway, as JAK1-STAT6 activation is controlled by small heterodimer partner-1 (SHP-1) (*PTPN6*) (46), and the JAK1-IRS2 pathway is controlled by *SOCS1* (47). Compatible with the DEG analysis of cluster 5, $CD23^+$/IL-$4R^+IgG^+$ memory B cells from highly sensitized peanut-allergic children had high proliferation and generated the most IgE cells after 7 days of culture compared to $CD23^+$/IL-$4R^+IgG^+$ memory B cells from peanut-allergic children with low sensitization or from non-allergic children. This suggests that in highly sensitized peanut-allergic children, higher concentrations of type 2 cytokines IL-4 and IL-13 may result in enhanced IL-4R-JAK1 signaling and class switching to IgE. JAK inhibition may thus directly suppress type 2-marked IgG^+ memory B cell activation and class switching to IgE. Importantly, a JAK inhibitor that is FDA-approved for the treatment of atopic dermatitis (48) is now being tested for the treatment of food allergy (clinical trial NCT05069831).

Several MHC II genes and genes related to the MHC II pathway were increased in the *IGHE*$^+$ cluster 5, suggesting that those cells had high competence for antigen presentation to CD4 T cells. Other DEGs with high significance were *HOPX* and *S100A10*. *HOPX* encodes an atypical homeobox protein that binds serum response factor (SRF) (49). *HOPX* function in B cells is not known but it was reported to be expressed in IgG^+CD27^+ B cells in the spleen (30) and in pre-plasmablasts of human tonsils (50); thus *HOPX* may mark high affinity switched memory cells prone to the plasma cell fate. The study by Koenig *et al.* (40) also identified an association between *IGHE* transcription and a type 2-marked cluster of memory B cells, referred by the authors as MBC2, characterized by *FCER2*,

IL4R, *IL13RA1*, and *HOPX* expression (40). They also identified a second similar cluster that lack *IGHE* expression. We did not identify a population of $IGHE^{neg}$ type 2-marked memory B cells, which may be explained by differences in patient populations as well as methodological analysis. The authors also described that about 25% of Ara h 2 tetramer binding B cells from the peripheral blood of peanut-allergic individuals expressed CD23 and IL-4R, whereas a much lower frequency of $CD23^{+}IL\text{-}4R^{+}$ cells were found among severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) receptor binding domain (RBD)-tetramer binding B cells in the same individuals (40). In this analysis, the Ara h 2 specificity was not corroborated by BCR sequencing and recombinant antibody expression.

Increased CD23 expression in B lymphocytes has been associated with allergic diseases and correlated with IgE production (23-25). It has been shown that IgE stabilizes the trimer of membrane CD23 on the cell surface and protects it from proteolysis (51). Thus, in addition to increased transcription, stabilization of CD23 by IgE in allergic patients may contribute to higher membrane expression. Interestingly, circulating IgE^{+} plasmablasts had higher expression of MHC II genes and *FCER2* than plasmablasts of other isotypes (6). Although this was considered evidence of lower maturation of IgE^{+} cells, it could reflect a developmental relationship between IgE^{+} cells and their potential precursors, $CD23^{+}IgG1^{+}$ memory B cells.

Previous repertoire analysis of circulating B cells found that the IgE repertoire was most closely related to IgG1, suggesting sequential class switching from IgG1 to IgE (10). In individuals undergoing sublingual immunotherapy (SLIT) for pollen allergy, clonotypes of IgG^{+} memory B cells that share BCR sequences with IgE^{+} plasma cells expressed *FCER2* and germline *IGHE* (11). A recent study described repertoire relatedness between IgA1 and IgE in gastrointestinal tissue from peanut-allergic patients, suggesting that IgA1 to IgE switching could take place in the gut (52). However, in our scRNA-seq analysis, very few $IgA1^{+}$ or IgM^{+} cells expressed *IGHE*, and *IGHE* transcription was strongly associated only with $IgG1^{+}$ and $IgG4^{+}$ memory B cells. It is also worth noting that antibodies derived from Ara h 2-binding IgM^{+} clones showed low or no somatic hypermutation and very low binding to Ara h 2.

Ara h 2-specific IgE has been reported to correlate with disease severity in peanut allergy (53). We found that the highly sensitized peanut-allergic children had a much higher frequency of Ara h 2-binding $IgG1^{+}$ memory B cells than children with low concentrations of peanut-specific IgE, again suggesting a relationship between peanut-specific $IgG1^{+}$ memory cells and peanut-specific IgE antibodies. Consistent with our results, other studies also found relatively high mutation rates in Ara h 2-specific antibody genes (6, 54, 55), and identified convergent sequence families of Ara h 2-binding BCRs, suggesting preferential binding of the Ara h 2 protein to selected germline VH and VL specificities that undergo further affinity maturation across individuals.

A limitation of the study was that repeated or multiple different analysis using the same blood sample could not be performed. This was due to the small blood volume collected, the fact that children have a lower frequency of memory B cells than adults, and that $CD23^{+}$ cells are a small percentage of memory cells. The scRNA-seq experiment was performed

with a relatively small number of samples and did not provide enough power for group comparisons. BCR repertoire relationships between the $CD23^+IgG^+$ memory B cells and the broad B cell compartment in each individual were not determined. This would have required a large BCR sequencing effort not possible to carry out within the timeline of this work.

In sum, we describe a population of somatically mutated $CD23^+$ *IGHE*$^+IgG1^+$ memory B cells with a high potential to respond to IL-4 and IL-13 and to undergo class switching to IgE. This population contains high affinity Ara h 2-specific memory B cells in highly sensitized peanut-allergic individuals. We propose that $CD23^+$ *IGHE*$^+IgG1^+$ memory B cells are involved in the persistence of food allergy by providing precursors of pathogenic IgE plasma cells. Thus, $CD23^+$ *IGHE*$^+IgG^+$ memory B cells are a potential therapeutic target in peanut allergy.

MATERIALS AND METHODS

Study design

This is a non-interventional study to identify memory B cell population related to persistent peanut allergy in children, we categorized pediatric peanut-allergic children with high peanut-specific IgE concentrations (over 100 kU_A/L) as possible persistent allergy population, pediatric peanut-allergic children with low peanut-specific IgE (less than 5 kU_A/L) as possible allergy-outgrowing population, and healthy children without any known allergy histories as non-allergic healthy population. The analysis was performed with PBMC samples from blood collected between 2019 and 2023. Peanut-allergic children with high concentrations of peanut-specific IgE (higher than 100 kU_A/L), low concentrations of peanut-specific IgE (lower than 5 kU_A/L) or intermediate concentrations of peanut-specific IgE (between 5 to 100 kU_A/L) were recruited among the pediatric population attending the Food Allergy clinic at Mount Sinai Hospitals. Peanut-allergic children recruited for the study had no history of previous immunotherapy or biological treatment for allergy, and did not suffer from immunodeficiency, autoimmunity, cancer, or any life-threatening infectious disease. Non-allergic children were recruited from the General Pediatrics clinic at the Mount Sinai Hospital. Non-allergic children were healthy children with no history of any atopic disease (e.g. food allergy, atopic dermatitis, asthma, rhinitis). Blood samples that yielded more than 30 x 10^6 PBMC were used for the study. Sample size justification, randomization and blinding were not performed. We hypothesized that $IgG1^+$ memory B cells are the precursors of high affinity, pathogenic IgE^+ plasma cells in persistent peanut allergy. Thus, we performed phenotypic and transcriptional analysis of total memory B cells by flow cytometry and 10X Genomics (WTA and BCR) scRNA-seq and characterized allergen-specific B cells by plate-based single-B cell analysis of BCR sequence and selected genes' expression. Functional analysis used in vitro cultures to identify the memory population that contained peanut-specific clones, and to determine the ability of memory cells to switch to IgE. The number of samples analyzed in each experiment is indicated in the figures and figure legends. Each experimental value was obtained from a unique sample, as only one blood sample was obtained from each donor.

Patient samples

For flow cytometry and the initial cell culture studies, PBMCs were isolated from peripheral blood from 58 pediatric patients with peanut allergy and 13 non-allergic children. Peanut-specific IgE (PsIgE) concentrations obtained on the day of blood draw were used to select the patients. Among 58 pediatric peanut-allergic patients, 18 patients had peanut-specific IgE higher than 100 kU_A/L, (PsIgE>100), 16 patients had sIgE lower than 5 kU_A/L (PsIgE<5), and 24 patients had sIgE between 8.0 to 76.6 kU_A/L (5<PsIgE<100). Non-allergic children were healthy children without any known history of malignancy, inflammatory or allergic disease, or any known infection at the time of blood collection. Four samples were obtained from PsIgE>100 patients for the additional cell culture studies. The age range and median were similar among cohorts: 5-14 years with a median age of 10 for peanut-allergic children with PsIgE>100; 4-15 years and a median age of 9 for peanut-allergic children with PsIgE<5; 6-17 years and a median age of 12 for peanut-allergic children with 5<PsIgE<100; 4-19 years and a median age of 11 for non-allergic children. In terms of incidence of other atopic diseases, 89% of peanut-allergic children with PsIgE>100 and 69% of peanut-allergic children with PsIgE<5 reported a history of AD or asthma, which was not significantly different between the two groups ($p=0.147$, table S1). All patients were screened by board-certified allergists/immunologists and were consented following ISMMS IRB-approved protocol.

Peripheral blood was collected in heparin tubes (BD, #367874) and processed within 24 hours. Blood was centrifuged at 1,500 rpm for 10 minutes. The plasma layer was collected and stored at −80°C. PBMCs were separated by Ficoll-Paque Plus density gradient centrifugation (Sigma-Aldrich, #GE17-1440-02) using a Leucosep tube (Grainer Bio-One, #227288) following the manufacturer's instructions. After the isolation, the PBMCs were counted and resuspended in 500 μl of 5% AB serum (GeminiBio, #100-52) in RPMI-1640 (Gibco, #11875-093) and 500 μl of 20% dimethyl sulfoxide (DMSO; Fisher, #BP-231-100) in AB serum at 0.5 to 10 x 10^6 PBMCs per cryovial and cryopreserved.

10X Genomics scRNA-seq of memory B cells

B cells were isolated from 17 million to 100 million PBMCs from 5 peanut-allergic children with sIgE>100 and 3 non-allergic children (table S2) by EasySep Human B cell Isolation Kit (STEMCELL technologies, #17954). Hash tagged antibodies were added to 3 million B cells at 1:50 from each sample. B cells from 2 peanut-allergic individuals were combined in experiment 1. In experiment 2, B cells from 3 peanut-allergic individuals were combined, and B cells from 3 non-allergic individuals were combined. The cells were resuspended in staining buffer consisting of Hanks' balanced salt solution (HBSS; Gibco, #14025-092) supplemented with 0.5 % bovine serum albumin (BSA; Fisher, #BP1600-1) and 1 mM ethylenediaminetetraacetic acid (EDTA; Invitrogen, #15575-038) at 10 to 20 x 10^6 cells/ml and stained for 20 minutes at 4°C with the following antibodies: Human TruStain FcX (1:20, human Fc receptor blocking solution, BioLegend, #422302, lot B313420), brilliant violet (BV) 421 anti-human CD19 (1:200, mouse clone HIB19, BioLegend, #302234, lot B275425) and fluorescein isothiocyanate (FITC) anti-human CD27 (1:200, mouse clone M-T271, BioLegend, #356404, lot B305082) antibodies and $CD19^+CD27^+$ memory B cell populations were sorted on a FACSAria II (BD Biosciences). After confirming cell number

and viability, B cells were run on a Chromium 10X controller using 5' chemistry (10X Genomics) in one lane (experiment 1) or two lanes (experiment 2) with an expected recovery rate of 10,000 cells per lane, according to the manufacturer's instructions. Libraries were generated and run on a HiSeq2500 apparatus.

10X Genomics scRNA-seq processing and analysis

10X Genomics 5' scRNA-seq data was processed with Cell Ranger version 3.1 and aligned to refdata-cellranger-GRCh38-3.0.0. Gene expression information from all individuals was processed using Seurat v4.1.1 (56) in R v4.1.0 (56). To remove apoptotic or lysing cells, cells with ≥10% of RNA transcripts from mitochondrial genes were excluded. To exclude poor quality cells, cells with reads from ≤ 400 features were also removed. To distinguish individuals within each sequencing run, hashtag oligonucleotide information was normalized using the centered log-ratio transformation. The individual of origin for each cell was determined using the default parameters of Seurat's HTODemux function. Only cells associated with a single individual were retained. Read counts were log-normalized using a scaling factor of 10^4. To account for variability in gene expression, log-normalized read counts were then scaled and centered for each feature. The top 2000 variable genes were then identified using Seurat's "vst" method. V, D, and J genes from the *IGH*, *IGL*, and *IGK* loci were removed so that the properties of the BCR expressed by the cell would remain independent of the cluster it is assigned to. Genes coding for the constant regions of antibody isotypes were also removed except for *IGHE*, which we intended to use as a marker gene. We then used Seurat's IntegrateData function to combine data from the three sequencing runs. Integration was performed using the previously identified top variable genes of each run, and the first 20 dimensions. Following integration, variable gene expression values were re-scaled and centered. This data was then reduced to the first 20 principal components.

B cell subtype identification utilized two clustering steps: initial clustering to identify and remove non-memory B cell contaminants, followed by secondary clustering to resolve memory B cell subtypes. For initial clustering, cells were clustered by Seurat's shared nearest neighbor clustering algorithm with a resolution of 0.25. The B cell subtype of each cluster was then determined by gene expression correlations to cell types in the immunoStates database (57). This resulted in 5 clusters identified as memory B cells and 2 clusters identified as T cells and naive B cells, which were removed. Within the remaining memory B cell clusters, data were re-scaled, reduced to the first 100 principal components, and clustered with a resolution of 0.5. This resulted in 10 memory B cell clusters, which were further confirmed to be memory B cells based on expression of marker genes (*CD24*, *TNFRSF13B*, fig. S6) and association with mutated BCRs (Fig. 2D).

To identify differentially expressed genes in each cluster, the FindAllMarkers function from Seurat was used. Non-integrated gene expression values in each cluster were compared to all other clusters combined. P values were calculated using a Wilcoxon rank-sum test and were adjusted using a Bonferroni adjustment. Genes were classified as differentially expressed within a cluster if their adjusted p values were < 0.05. To identify the enrichment of gene ontologies for each cluster, the list of positively differentially expressed genes for

each cluster was tested using *enrichR* v3.0 (58, 59). Differentially expressed genes were compared against the GO Biological Process 2021 database (58, 59) within the *enrichR* framework (60).

PacBio single cell BCR sequencing of Ara h 2-binding cells

Ara h 2 multimers were generated by associating biotinylated Ara h 2 with phycoerythrin (PE)-labeled streptavidin (Invitrogen, #S866) or Alexa Fluor (AF647)-labeled streptavidin (Invitrogen, #S32357). Similarly, detoxified diphtheria toxin (DT) (CRM197, PFENEX Biopharmaceuticals) was used to generate DT multimers. Pre-enriched B cells by EasySep Human B cell Isolation Kit (STEMCELL technologies, #17954) from 8 peanut-allergic individuals with PsIgE>100, 5 peanut-allergic individuals with PsIgE<5, and 2 non-allergic individuals were used for Ara h 2 binder or DT binder sorting (table S5). Enriched B cells were resuspended in staining buffer consisting of HBSS (Gibco, #14025-092) supplemented with 0.5 % BSA (Fisher, #BP1600-1) and 1 mM EDTA (Invitrogen, #15575-038) at 10 to 20 x 10^6 cells/ml and stained for 20 minutes at 4°C with the following antibodies: Human TruStain FcX (1:20, human Fc receptor blocking solution, BioLegend, #422302, lot B313420), Alexa Fluor 647 anti-human CD19 (1:200, mouse clone HIB19, BioLegend, #302220, lot B260236), Ara h 2-PE (1:200), and Ara h 2-APC (1:200) or DT-PE (1:200) and DT-APC (1:200). Single Ara h 2-binding or DT-binding B cell was sorted using a FACSAria II (BD Biosciences) into each well in the 96 well plates with lysis buffer (Takara, # 635013) containing RNAse inhibitor (Promega, #N261B), dNTPs (New England Biolabs, NEB #N0447), and oligo dT-VN30 anchored with template switch oligo sequence (table S6). Plates were flash-frozen in dry ice and stored at −80°C. Plates were thawed on ice, then oligo dT was added and annealing was performed at 70°C for 5 minutes. cDNA synthesis and cDNA amplification was performed with template switching oligo (TSO) using template switching RT enzyme mix (NEB, #M0466) according to the manufacturer's instruction. The first PCR was performed to amplify BCR heavy and light chains using outer primers targeting human immunoglobulin constant regions (IgG, IgM for heavy chains, kappa, and lambda for light chains) and TSO inner primer (table S6), respectively. After cleaning up the PCR products, BCR heavy and light chains were further amplified with inner primers with plate and row-specific ID (5') and column-specific ID (3') to enable the identification of BCR sequences derived from each well (table S6). Afterward, all PCR products were combined as one library. PacBio sequencing was performed to obtain long amplicon reads to cover 500 to 1000 bp. All PCR reactions were performed using the NEBNext Ultra II Q5 master mix (NEB, #M0544).

B cell receptor sequence processing and analysis

BCR sequencing was obtained from two sources: 10X Genomics scRNA-seq + BCR sequencing and PacBio single cell BCR sequencing. BCRs from both were processed using the Immcantation suite (immcantation.org). BCR sequence data from 10X Genomics scRNA-seq + BCR sequencing began with the filtered V(D)J contigs from 10X Genomics Cell Ranger version 3.1. To obtain V and J gene assignments, these contigs were aligned to the IMGT GENE-DB germline reference allele database (obtained 8/3/2019) (61). Non-productive heavy and light chain BCR sequences were removed. BCRs were annotated by B cell subtype based on matching their single cell barcodes to the gene expression information.

Only cells that had both gene expression and BCR sequence data were retained. In cells with multiple heavy chains, only one heavy chain with the highest unique molecular identifier count was retained. In the event of a tie, the first heavy chain identified was retained.

For BCR sequencing data from sorted single cell PacBio sequencing data, adaptor ligation, sequencing reactions, and initial sequence processing were conducted at GENEWIZ, LLC/Azenta US, Inc using standard PacBio circular consensus sequence (CCS) tools and guidelines. Further processing was performed using *presto* v0.6.2 (62). To remove reads too short to likely contain functional BCRs and associated constant regions, reads shorter than 500bp were removed. To identify the plate and row of each read, forward primers were aligned to the first 1000bp of each read, with upstream sequence sites removed. To identify *IGHG* sub-isotype sequences, internal constant region sequences associated with each subisotype were aligned to each read. Reads were tagged if they contained a match to an *IGHG* sub-isotype constant region with a maximum error rate of 0.1. To identify plate column and constant region primer (IgG, IgM, IgK, IgL) for each read, reverse primer sequences were also aligned to each read. Primer sequences are available in table S6. Nucleotide sites matching primer sequences, as well as the sites downstream of them, were removed. Reads failing to exactly match both a forward and reverse primer were discarded. Remaining reads shorter than 500bp in length were also discarded. After initial processing, PacBio BCR sequence reads were assembled into consensus sequences. Because each well within each plate should represent reads from a single B cell, reads within a forward and reverse primer combination should correspond to a single consensus sequence. However, initial analyses indicated further filtering was required to obtain consensus sequences with appropriate confidence. Within each forward and reverse primer combination, reads were further clustered using *vsearch* v2.14.1 (63) with an 80% similarity threshold. To account for indel variation within a cluster, reads within these clusters were aligned to each other using *MUSCLE* v3.8.1551 (64). Reads within each of these clusters were assembled into consensus sequences. Nucleotide sites in which < 90% of reads had the same character, or had a consensus quality score < 20, were assigned as ambiguous “N” characters. Sites in which >50% of reads contained gap characters were removed. Read groups containing less than 10 reads or in which >10% of sites differed from their final consensus sequence were removed. As with 10X Genomics BCR sequence data, consensus sequences were aligned to the IMGT GENE-DB database (61) to obtain V and J germline annotations. Non-productive heavy and light chain BCR sequences were removed. 18.3% of forward and reverse primer combinations yielded multiple productive consensus sequences from different clusters. These sequences possibly represent doublets and were discarded unless one sequence, which was retained, accounted for at least 90% of the reads in that well or constant region group. For each consensus sequence, *IGHG* sub-isotype was determined as the most frequent-matching sub-isotype sequence among constituent reads, or “Unknown” if no sub-isotype sequence match was found. Confirming the accuracy of this step, *IGHG* sub-isotypes could only be determined for sequences that also matched *IGHG* primers. As a further quality control, 16 sequences originated from wells in which no Ig DNA was detected through PCR and were discarded. Additionally, 13 sequences with < 300 non-ambiguous nucleotides were also discarded. This pipeline yielded 243 total heavy chain sequences, 210 of which were paired with functional light chain BCRs. We validated our

consensus sequences through two tests. First, consensus sequences assembled with high error rates are unlikely to exactly match IMGT reference sequences. We tested this in our data by quantifying the SHM frequency of *IGHM* heavy chain consensus sequences. Although sequences from *IGHG* primers were highly mutated (median SHM = 8.7%), most *IGHM* consensus sequences were mostly unmutated (median SHM = 0%, Fig. 5A). Next, heavy and light chain sequences assembled from different B cells are unlikely to have similar frequencies of SHM. In our consensus sequences, however, SHM frequencies of paired heavy and light chains were closely correlated (linear regression slope $p<2e-16$, $R^2=0.63$).

For BCR data from both data sources, SHM frequencies was determined for each cell as the frequency of non-ambiguous mismatches along the IGHV gene of the unmutated germline sequence along IMGT positions 1-312 using *shazam* v1.1.2 (65). To identify potential Ara h 2-binding convergent sequence families across all single cell datasets, sequences from all datasets were partitioned based on common *IGHV* and *IGHJ* gene annotations, as well as junction region length. Within these groups, sequences differing from one another by an amino acid Hamming distance threshold of 0.2 within the junction region were clustered together using single linkage hierarchical clustering (66) implemented in *scoper* v1.2.0 (67). When searching against public sequences, we also identified one cell (P4AC5) with a matching V gene, J gene, and junction amino acid Hamming distance of 0.21 compared to P413P1E10. This cell was also included within the P413P1E10 family and confirmed to bind to Ara h 2 through ELISA (Fig. 6B). In the 4C5G1 convergent sequence family, all cells expressed similar IGK light chains, however, cell P1DC5 also expressed an IGL light chain (IGLV4-69*01/IGLJ3*02). This sequence was not included in further analyses. Consensus unmutated V and J gene segments for each convergent sequence family (Fig. 6A) were constructed using the createGermlines function in *dowser* v1.1.0 (68). For ambiguous V gene assignments, only the first assignment was used for reconstruction.

Statistical analysis

All raw, individual-level data for experiments where $n<20$ are presented in data file S4. Unless otherwise specified, data were analyzed and graphed with GraphPad Prism 9 software (GraphPad Software). The Shapiro-Wilk test was performed to select either one-way ANOVA or the Kruskal-Wallis test for the statistical significance for >2 groups. Either the two-sided student’s t-test or two-sided Mann-Whitney U test were used to perform parametric or nonparametric test, respectively, for two group comparisons with 95 % confidential interval for median. For correlation analysis, Shapiro-Wilk test was performed to select a nonparametric Spearman correlation test. The Chi-square test was used to determine the significant differences between expected frequencies and observed frequencies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank Jiaming Lin, Charuta Aguti, and clinical personnel of the Jaffe Food Allergy Institute and General Pediatrics at ISMMS for patient recruitment; Ankita Prakash and Pedro Silva for general lab support; the Human Immuno Monitoring Center, and the Genomics and Flow Cytometry Cores at ISMMS for technical support; James Duty for help with Octet experiments, Deepta Bhattacharya for hCD40LB21 cells, and Juan J. Lafaille, Goran Bajic, and members of the MACL laboratory for critical comments.

Funding:

This work was supported by National Institutes of Health (NIH) grants R01AI151707 (to MACL), R01 AI153708 (to MACL), R01AI104739 (to SHK), and K99AI159302 (to KBH). MO received a fellowship from the ISMMS T32 NIH training program (AI078892) and JR was awarded an NIH pre-doctoral fellowship (F31AI172404).

References and Notes

1. Sicherer SH, Sampson HA, Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. J Allergy Clin Immunol 141, 41–58 (2018). [PubMed: 29157945]
2. Xiong H, Dolpady J, Wabl M, Curotto de Lafaille MA, Lafaille JJ, Sequential class switching is required for the generation of high affinity IgE antibodies. J Exp Med 209, 353–364 (2012). [PubMed: 22249450]
3. He JS, Subramaniam S, Narang V, Srinivasan K, Saunders SP, Carbajo D, Wen-Shan T, Hidayah Hamadee N, Lum J, Lee A, Chen J, Poidinger M, Zolezzi F, Lafaille JJ, Curotto de Lafaille MA, IgG1 memory B cells keep the memory of IgE responses. Nat Commun 8, 641 (2017). [PubMed: 28935935]
4. Berkowska MA, Heeringa JJ, Hajdarbegovic E, van der Burg M, Thio HB, van Hagen PM, Boon L, Orfao A, van Dongen JJ, van Zelm MC, Human IgE(+) B cells are derived from T cell-dependent and T cell-independent pathways. J Allergy Clin Immunol 134, 688–697 e686 (2014). [PubMed: 24835500]
5. Heeringa JJ, Rijvers L, Arends NJ, Driessen GJ, Pasmans SG, van Dongen JJM, de Jongste JC, van Zelm MC, IgE-expressing memory B cells and plasmablasts are increased in blood of children with asthma, food allergy, and atopic dermatitis. Allergy, (2018).
6. Croote D, Darmanis S, Nadeau KC, Quake SR, High-affinity allergen-specific human antibodies cloned from single IgE B cell transcriptomes. Science 362, 1306–1309 (2018). [PubMed: 30545888]
7. Saunders SP, Ma EGM, Aranda CJ, Curotto de Lafaille MA, Non-classical B Cell Memory of Allergic IgE Responses. Front Immunol 10, 715 (2019). [PubMed: 31105687]
8. Miranda-Waldetario MCG, Redes J, Fernandes-Braga W, Curotto de Lafaille MA, B cell memory of IgE responses in food allergy. *Reference Module in Food Allergy* **Encyclopedia of Food Allergy**, (2023).
9. Takhar P, Corrigan CJ, Smurthwaite L, O'Connor BJ, Durham SR, Lee TH, Gould HJ, Class switch recombination to IgE in the bronchial mucosa of atopic and nonatopic patients with asthma. J Allergy Clin Immunol 119, 213–218 (2007). [PubMed: 17208604]
10. Looney TJ, Lee JY, Roskin KM, Hoh RA, King J, Glanville J, Liu Y, Pham TD, Dekker CL, Davis MM, Boyd SD, Human B-cell isotype switching origins of IgE. J Allergy Clin Immunol 137, 579–586.e577 (2016). [PubMed: 26309181]
11. Hoof I, Schulten V, Layhadi JA, Stranzl T, Christensen LH, de la Mata S. Herrera, Seumois G, Vijayanand P, Lundegaard C, Niss K, Lund A, Ahrenfeldt J, Holm J, Steveling E, Sharif H, Durham SR, Peters B, Shamji MH, Andersen PS, Allergen-specific IgG(+) memory B cells are temporally linked to IgE memory responses. J Allergy Clin Immunol 146, 180–191 (2020). [PubMed: 31883847]
12. Hofmaier S, Hatzler L, Rohrbach A, Panetta V, Hakimeh D, Bauer CP, Hoffman U, Forster J, Zepp F, Schuster A, Stock P, Wahn U, Keil T, Lau S, Matricardi PM, "Default" versus "pre-atopic" IgG responses to foodborne and airborne pathogenesis-related group 10 protein molecules in

birch-sensitized and nonatopic children. J Allergy Clin Immunol 135, 1367–1374.e1361-1368 (2015). [PubMed: 25458000]

13. Schwarz A, Panetta V, Cappella A, Hofmaier S, Hatzler L, Rohrbach A, Tsilochristou O, Bauer CP, Hoffmann U, Forster J, Zepp F, Schuster A, D'Amelio R, Wahn U, Keil T, Lau S, Matricardi PM, IgG and IgG. J Allergy Clin Immunol 138, 1426–1433.e1412 (2016). [PubMed: 27264457]
14. Neuman-Sunshine DL, Eckman JA, Keet CA, Matsui EC, Peng RD, Lenehan PJ, Wood RA, The natural history of persistent peanut allergy. Ann Allergy Asthma Immunol 108, 326–331.e323 (2012). [PubMed: 22541403]
15. Sicherer SH, Wood RA, Vickery BP, Jones SM, Liu AH, Fleischer DM, Dawson P, Mayer L, Burks AW, Grishin A, Stablein D, Sampson HA, The natural history of egg allergy in an observational cohort. J Allergy Clin Immunol 133, 492–499 (2014). [PubMed: 24636473]
16. Wood RA, Sicherer SH, Vickery BP, Jones SM, Liu AH, Fleischer DM, Henning AK, Mayer L, Burks AW, Grishin A, Stablein D, Sampson HA, The natural history of milk allergy in an observational cohort. J Allergy Clin Immunol 131, 805–812 (2013). [PubMed: 23273958]
17. Savage J, Sicherer S, Wood R, The Natural History of Food Allergy. J Allergy Clin Immunol Pract 4, 196–203; quiz 204 (2016). [PubMed: 26968958]
18. Sicherer SH, Sampson HA, 9. Food allergy. J Allergy Clin Immunol 117, S470–475 (2006). [PubMed: 16455349]
19. Skolnick HS, Conover-Walker MK, Koerner CB, Sampson HA, Burks W, Wood RA, The natural history of peanut allergy. J Allergy Clin Immunol 107, 367–374 (2001). [PubMed: 11174206]
20. Fleischer DM, Conover-Walker MK, Christie L, Burks AW, Wood RA, The natural progression of peanut allergy: Resolution and the possibility of recurrence. J Allergy Clin Immunol 112, 183–189 (2003). [PubMed: 12847497]
21. Santos AF, Du Toit G, O'Rourke C, Becares N, Couto-Francisco N, Radulovic S, Khaleva E, Basting M, Harris KM, Larson D, Sayre P, Plaut M, Roberts G, Bahnson HT, Lack G, Biomarkers of severity and threshold of allergic reactions during oral peanut challenges. J Allergy Clin Immunol 146, 344–355 (2020). [PubMed: 32311390]
22. Foong RX, Santos AF, Biomarkers of diagnosis and resolution of food allergy. Pediatr Allergy Immunol 32, 223–233 (2021). [PubMed: 33020989]
23. Czarnowicki T, Esaki H, Gonzalez J, Renert-Yuval Y, Brunner P, Oliva M, Estrada Y, Xu H, Zheng X, Talasila S, Haugh I, Huynh T, Lyon S, Tran G, Sampson H, Suárez-Fariñas M, Krueger JG, Guttman-Yassky E, Paller AS, Alterations in B-cell subsets in pediatric patients with early atopic dermatitis. J Allergy Clin Immunol 140, 134–144.e139 (2017). [PubMed: 27965110]
24. Chary AV, Hemalatha R, Murali MV, Jayaprakash D, Kumar BD, Association of T-regulatory cells and CD23/CD21 expression with vitamin D in children with asthma. Ann Allergy Asthma Immunol 116, 447–454.e442 (2016). [PubMed: 27026514]
25. Selb R, Eckl-Dorna J, Neunkirchner A, Schmetterer K, Marth K, Gamper J, Jahn-Schmid B, Pickl WF, Valenta R, Niederberger V, CD23 surface density on B cells is associated with IgE levels and determines IgE-facilitated allergen uptake, as well as activation of allergen-specific T cells. J Allergy Clin Immunol 139, 290–299.e294 (2017). [PubMed: 27372566]
26. Klein U, Rajewsky K, Küppers R, Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. J Exp Med 188, 1679–1689 (1998). [PubMed: 9802980]
27. Sutton HJ, Aye R, Idris AH, Vistein R, Nduati E, Kai O, Mwacharo J, Li X, Gao X, Andrews TD, Koutsakos M, Nguyen THO, Nekrasov M, Milburn P, Eltahla A, Berry AA, Kc N, Chakravarty S, Sim BKL, Wheatley AK, Kent SJ, Hoffman SL, Lyke KE, Bejon P, Luciani F, Kedzierska K, Seder RA, Ndungu FM, Cockburn IA, Atypical B cells are part of an alternative lineage of B cells that participates in responses to vaccination and infection in humans. Cell Rep 34, 108684 (2021). [PubMed: 33567273]
28. Glass DR, Tsai AG, Oliveria JP, Hartmann FJ, Kimmey SC, Calderon AA, Borges L, Glass MC, Wagar LE, Davis MM, Bendall SC, An Integrated Multi-omic Single-Cell Atlas of Human B Cell Identity. Immunity 53, 217–232 e215 (2020). [PubMed: 32668225]

29. Aranda CJ, Gonzalez-Kozlova E, Saunders SP, Fernandes-Braga W, Ota M, Narayanan S, Jin-Shu H, Duca E, Bose S, Gnjatic S, Shattner G, Reibman J, Soter NA, Guttman-Yassky E, Lafaille MAC, IgG memory B cells expressing IL4R and FCER2 are associated with atopic diseases. Allergy, (2022).
30. Descatoire M, Weller S, Irtan S, Sarnacki S, Feuillard J, Storck S, Guiochon-Mantel A, Bouligand J, Morali A, Cohen J, Jacquemin E, Iascone M, Bole-Feysot C, Cagnard N, Weill JC, Reynaud CA, Identification of a human splenic marginal zone B cell precursor with NOTCH2-dependent differentiation properties. J Exp Med 211, 987–1000 (2014). [PubMed: 24733829]
31. Seo JS, Svenningsson P, Modulation of Ion Channels and Receptors by p11 (S100A10). Trends Pharmacol Sci 41, 487–497 (2020). [PubMed: 32418644]
32. Rothman P, Lutzker S, Cook W, Coffman R, Alt FW, Mitogen plus interleukin 4 induction of C epsilon transcripts in B lymphoid cells. J Exp Med 168, 2385–2389 (1988). [PubMed: 3264328]
33. Geha RS, Jabara HH, Brodeur SR, The regulation of immunoglobulin E class-switch recombination. Nat Rev Immunol 3, 721–732 (2003). [PubMed: 12949496]
34. He JS, Meyer-Hermann M, Xiangying D, Zuan LY, Jones LA, Ramakrishna L, de Vries VC, Dolpady J, Aina H, Joseph S, Narayanan S, Subramaniam S, Puthia M, Wong G, Xiong H, Poidinger M, Urban JF, Lafaille JJ, Curotto de Lafaille MA, The distinctive germinal center phase of IgE+ B lymphocytes limits their contribution to the classical memory response. J Exp Med 210, 2755–2771 (2013). [PubMed: 24218137]
35. Dang TD, Tang M, Choo S, Licciardi PV, Koplin JJ, Martin PE, Tan T, Gurrin LC, Ponsonby AL, Tey D, Robinson M, Dharmage SC, Allen KJ, study H, Increasing the accuracy of peanut allergy diagnosis by using Ara h 2. J Allergy Clin Immunol 129, 1056–1063 (2012). [PubMed: 22385632]
36. Klemans RJ, Broekman HC, Knol EF, Bruijnzeel-Koomen CA, Otten HG, Pasmans SG, Knulst AC, Ara h 2 is the best predictor for peanut allergy in adults. J Allergy Clin Immunol Pract 1, 632–638.e631 (2013). [PubMed: 24565711]
37. Hemmings O, Du Toit G, Radulovic S, Lack G, Santos AF, Ara h 2 is the dominant peanut allergen despite similarities with Ara h 6. J Allergy Clin Immunol 146, 621–630.e625 (2020). [PubMed: 32298698]
38. Suprun M, Sicherer SH, Wood RA, Jones SM, Leung DYM, Henning AK, Dawson P, Burks AW, Lindblad R, Getts R, Suarez-Farinas M, Sampson HA, Early epitope-specific IgE antibodies are predictive of childhood peanut allergy. J Allergy Clin Immunol 146, 1080–1088 (2020). [PubMed: 32795587]
39. Suprun M, Kearney P, Hayward C, Butler H, Getts R, Sicherer SH, Turner PJ, Campbell DE, Sampson HA, Predicting probability of tolerating discrete amounts of peanut protein in allergic children using epitope-specific IgE antibody profiling. Allergy 77, 3061–3069 (2022). [PubMed: 35960650]
40. Joshua NPHK Koenig FE, Phelps Allyssa, Bruton Kelly, Hoof Ilka, Lund Gitte, Libera Danielle Della, Lund Anders, Christensen Lars Harder, Glass David R., Walker Tina, Fang Allison, Waserman Susan, Jordana Manel, Andersen Peter S, A Distinct Phenotype of Polarized Memory B cell holds IgE Memory. **DOI:** 10.1126/scitranslmed.adi0944, (2024).**DOI:**
41. Stewart A, Ng JC, Wallis G, Tsioligka V, Fraternali F, Dunn-Walters DK, Single-Cell Transcriptomic Analyses Define Distinct Peripheral B Cell Subsets and Discrete Development Pathways. Front Immunol 12, 602539 (2021). [PubMed: 33815362]
42. McKenzie CI, Varese N, Aui PM, Reinwald S, Wines BD, Hogarth PM, Thien F, Hew M, Rolland JM, O'Hehir RE, van Zelm MC, RNA sequencing of single allergen-specific memory B cells after grass pollen immunotherapy: Two unique cell fates and CD29 as a biomarker for treatment effect. Allergy 78, 822–835 (2023). [PubMed: 36153670]
43. Erazo A, Kutchukhidze N, Leung M, Christ AP, Urban JF Jr., de Lafaille M. A. Curotto, Lafaille JJ, Unique maturation program of the IgE response in vivo. Immunity 26, 191–203 (2007). [PubMed: 17292640]
44. Zamorano J, Kelly AE, Austrian J, Wang HY, Keegan AD, Costimulation of resting B lymphocytes alters the IL-4-activated IRS2 signaling pathway in a STAT6 independent manner: implications for cell survival and proliferation. Cell Res 11, 44–54 (2001). [PubMed: 11305324]

45. McCormick SM, Heller NM, Commentary: IL-4 and IL-13 receptors and signaling. Cytokine 75, 38–50 (2015). [PubMed: 26187331]
46. Hanson EM, Dickensheets H, Qu CK, Donnelly RP, Keegan AD, Regulation of the dephosphorylation of Stat6. Participation of Tyr-713 in the interleukin-4 receptor alpha, the tyrosine phosphatase SHP-1, and the proteasome. J Biol Chem 278, 3903–3911 (2003). [PubMed: 12459556]
47. McCormick SM, Gowda N, Fang JX, Heller NM, Suppressor of Cytokine Signaling (SOCS)1 Regulates Interleukin-4 (IL-4)-activated Insulin Receptor Substrate (IRS)-2 Tyrosine Phosphorylation in Monocytes and Macrophages via the Proteasome. J Biol Chem 291, 20574–20587 (2016). [PubMed: 27507812]
48. Deeks ED, Duggan S, Abrocitinib: First Approval. Drugs 81, 2149–2157 (2021). [PubMed: 34807428]
49. Liu Y, Zhang W, The role of HOPX in normal tissues and tumor progression. Biosci Rep 40, (2020).
50. King HW, Orban N, Riches JC, Clear AJ, Warnes G, Teichmann SA, James LK, Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. Sci Immunol 6, (2021).
51. Hibbert RG, Teriete P, Grundy GJ, Beavil RL, Reljic R, Holers VM, Hannan JP, Sutton BJ, Gould HJ, McDonnell JM, The structure of human CD23 and its interactions with IgE and CD21. J Exp Med 202, 751–760 (2005). [PubMed: 16172256]
52. Hoh RA, Joshi SA, Lee JY, Martin BA, Varma S, Kwok S, Nielsen SCA, Nejad P, Haraguchi E, Dixit PS, Shutthanandan SV, Roskin KM, Zhang W, Tupa D, Bunning BJ, Manohar M, Tibshirani R, Fernandez-Becker NQ, Kambham N, West RB, Hamilton RG, Tsai M, Galli SJ, Chinthrajah RS, Nadeau KC, Boyd SD, Origins and clonal convergence of gastrointestinal IgE. Sci Immunol 5, (2020).
53. Martinet J, Couderc L, Renosi F, Bobée V, Marguet C, Boyer O, Diagnostic Value of Antigen-Specific Immunoglobulin E Immunoassays against Ara h 2 and Ara h 8 Peanut Components in Child Food Allergy. Int Arch Allergy Immunol 169, 216–222 (2016). [PubMed: 27225199]
54. Hoh RA, Joshi SA, Liu Y, Wang C, Roskin KM, Lee JY, Pham T, Looney TJ, Jackson KJL, Dixit VP, King J, Lyu SC, Jenks J, Hamilton RG, Nadeau KC, Boyd SD, Single B-cell deconvolution of peanut-specific antibody responses in allergic patients. J Allergy Clin Immunol 137, 157–167 (2016). [PubMed: 26152318]
55. Patil SU, Ogunniyi AO, Calatroni A, Tadigotla VR, Ruiter B, Ma A, Moon J, Love JC, Shreffler WG, Peanut oral immunotherapy transiently expands circulating Ara h 2-specific B cells with a homologous repertoire in unrelated subjects. J Allergy Clin Immunol 136, 125–134.e112 (2015). [PubMed: 25985925]
56. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LM, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P, Satija R, Integrated analysis of multimodal single-cell data. Cell 184, 3573–3587.e3529 (2021). [PubMed: 34062119]
57. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R, Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36, 411–420 (2018). [PubMed: 29608179]
58. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25, 25–29 (2000). [PubMed: 10802651]
59. G. O. Consortium, The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res 49, D325–D334 (2021). [PubMed: 33290552]
60. Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, Lachmann A, Wojciechowicz ML, Kropiwnicki E, Jagodnik KM, Jeon M, Ma'ayan A, Gene Set Knowledge Discovery with Enrichr. Curr Protoc 1, e90 (2021). [PubMed: 33780170]

61. Giudicelli V, Chaume D, Lefranc MP, IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. Nucleic Acids Res 33, D256–261 (2005). [PubMed: 15608191]

62. Vander Heiden JA, Yaari G, Uduman M, Stern JN, O'Connor KC, Hafler DA, Vigneault F, Kleinstein SH, pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. Bioinformatics 30, 1930–1932 (2014). [PubMed: 24618469]

63. Rognes T, Flouri T, Nichols B, Quince C, Mahé F, VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584 (2016). [PubMed: 27781170]

64. Edgar RC, MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792–1797 (2004). [PubMed: 15034147]

65. Yaari G, Vander Heiden JA, Uduman M, Gadala-Maria D, Gupta N, Stern JN, O'Connor KC, Hafler DA, Laserson U, Vigneault F, Kleinstein SH, Models of somatic hypermutation targeting and substitution based on synonymous mutations from high-throughput immunoglobulin sequencing data. Front Immunol 4, 358 (2013). [PubMed: 24298272]

66. Gupta NT, Adams KD, Briggs AW, Timberlake SC, Vigneault F, Kleinstein SH, Hierarchical Clustering Can Identify B Cell Clones with High Confidence in Ig Repertoire Sequencing Data. J Immunol 198, 2489–2499 (2017). [PubMed: 28179494]

67. Nouri N, Kleinstein SH, A spectral clustering-based method for identifying clones from high-throughput B cell repertoire sequencing data. Bioinformatics 34, i341–i349 (2018). [PubMed: 29949968]

68. Hoehn KB, Pybus OG, Kleinstein SH, Phylogenetic analysis of migration, differentiation, and class switching in B cells. PLoS Comput Biol 18, e1009885 (2022). [PubMed: 35468128]

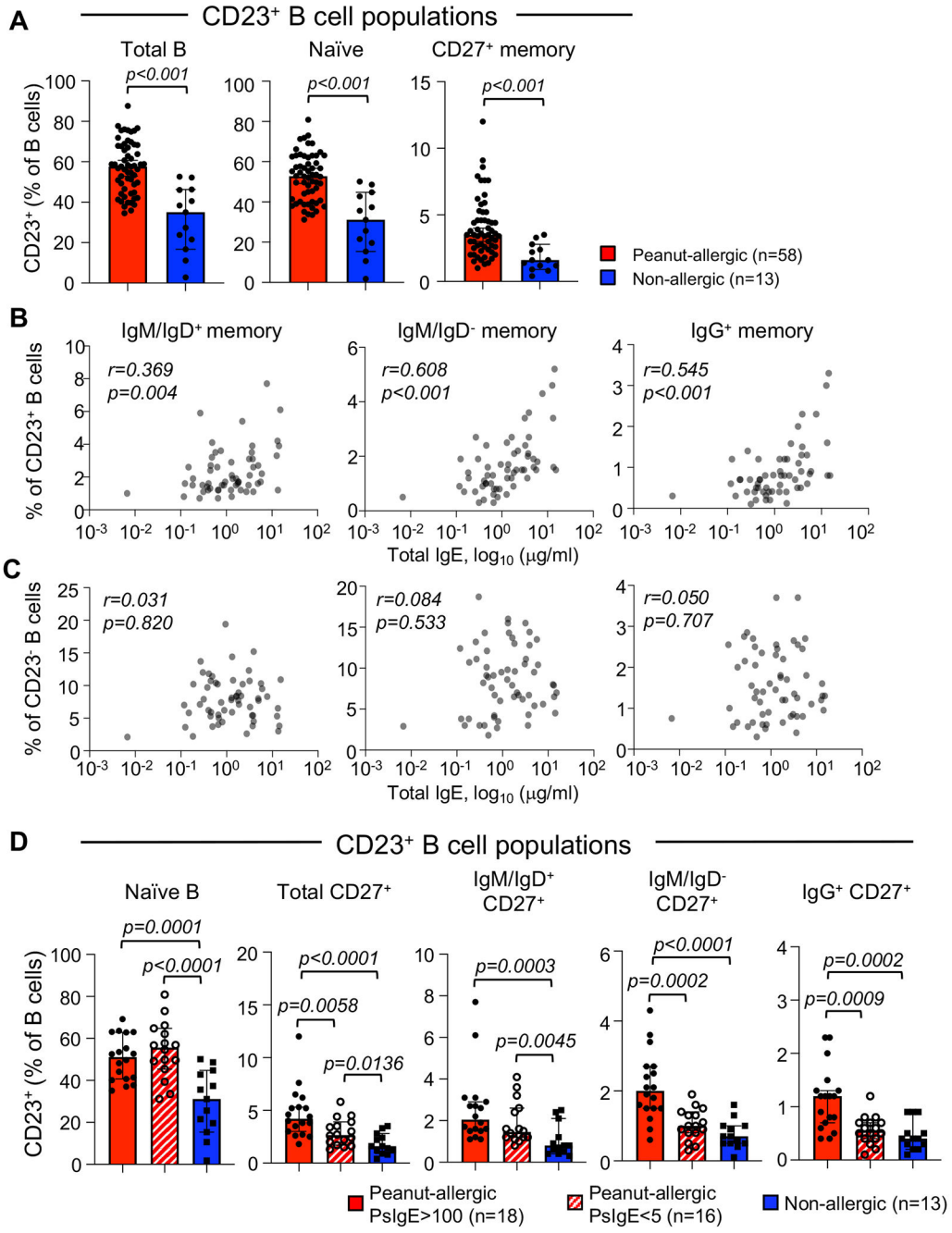

Fig. 1. Frequency of IgG$^+$CD23$^+$ memory B cells correlates with total IgE concentrations in peanut allergy.

(**A**) Percentages of surface CD23$^+$ B cells in total CD19$^+$ B cells, in CD19$^+$CD27$^-$IgM/IgD$^+$ B cells (Naïve), and in CD19$^+$CD27$^+$ B cells (CD27$^+$ memory) were analyzed in peanut-allergic patients (n=58; red dotted bar) and non-allergic healthy children (n=13; blue solid bar). Data are presented as median ± 95 % CI. Data were analyzed by unpaired t-test. (**B and C**) Shown are correlations between the frequency of CD23$^+$ memory B cells (B) or CD23$^-$ memory B cells (C) with total plasma IgE concentrations. Data were analyzed by Spearman correlation. (**D**) CD23$^+$ B cell populations were analyzed among

peanut-allergic patients with high peanut-specific IgE (PA PsIgE>100, n=18; red solid bar) or with low peanut-specific IgE (PA PsIgE<5, n=16; red shaded bar), as well as in non-allergic individuals (n=13; blue solid bar). Data are presented as median ± 95 % CI. Data were analyzed by unpaired t-test for naïve B and switched $CD27^+$ or Mann-Whitney test for total $CD27^+$ and IgG^+ $CD27^+$.

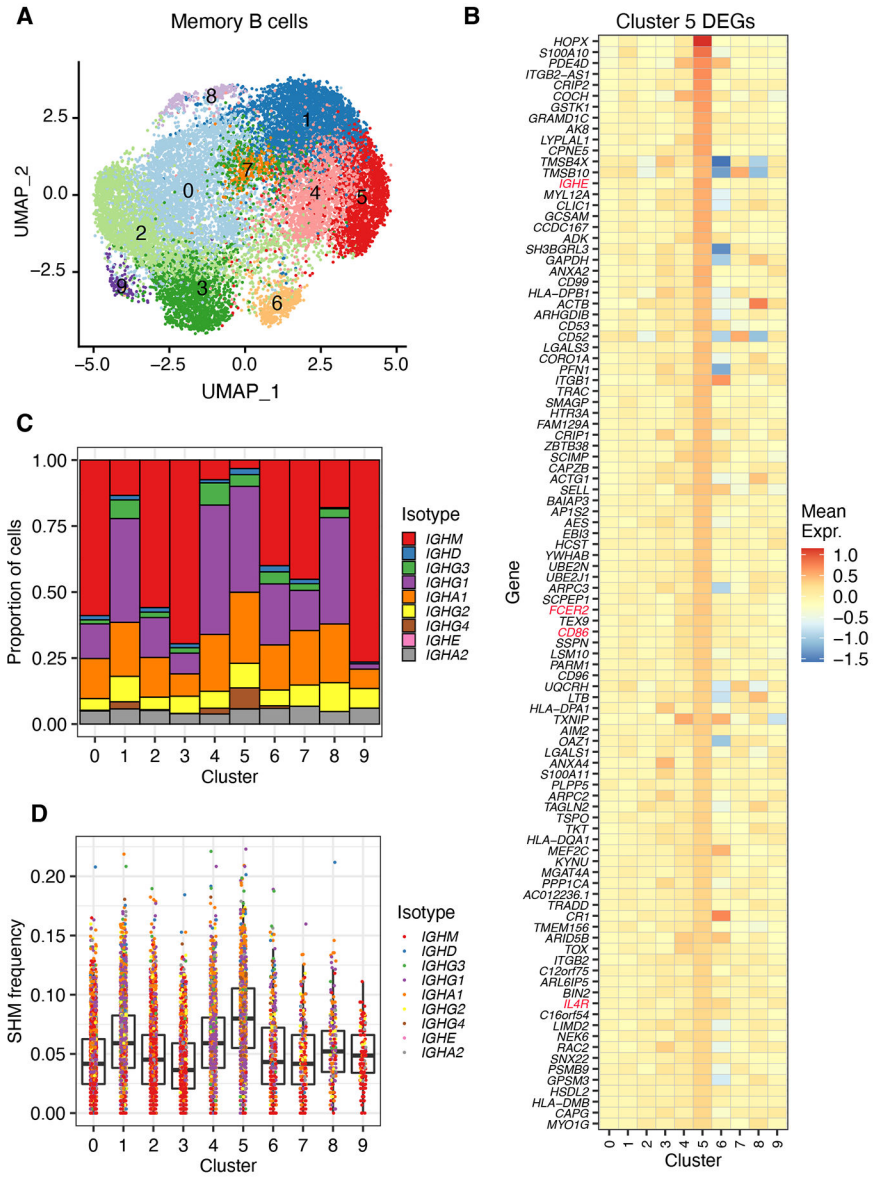

Fig. 2. scRNA-seq seq reveals CD23^{+} memory B cells transcribe *IGHE*, are enriched for *IGHG* BCR constant regions, and are highly mutated.

(**A**) Shown is a Uniform Manifold Approximation and Projection (UMAP) plot of cells obtained from peanut-allergic and non-allergic individuals. Each point represents a single cell. (**B**) Shown is scaled gene expression of cells in cluster 5. Tiles show scaled, normalized expression of cluster 5. Genes are positively differentially expressed genes (adjusted p < 0.05) for cluster 5. IL-4 and IL-13 regulated genes *IGHE*, *FCER2*, *IL4R*, and *CD86* are shown in red. (**C**) Shown is the proportion of cells in each scRNA-seq defined cluster that uses each isotype constant region. Enrichment for *IGHG*-associated BCRs was analyzed by chi-square test (Cluster 5: 61.7% of cells, $p < 2.2e\text{-}16$). (**D**) Shown is the frequency of IGHV-gene SHM for cells in each cluster. Cells are colored by associated constant region. Data in (D) are presented as box and whiskers plots in which boxes represent the median as

well as first and third quartiles, and whiskers extend to the largest and smallest values within 1.5 times the interquartile range from the boxes.

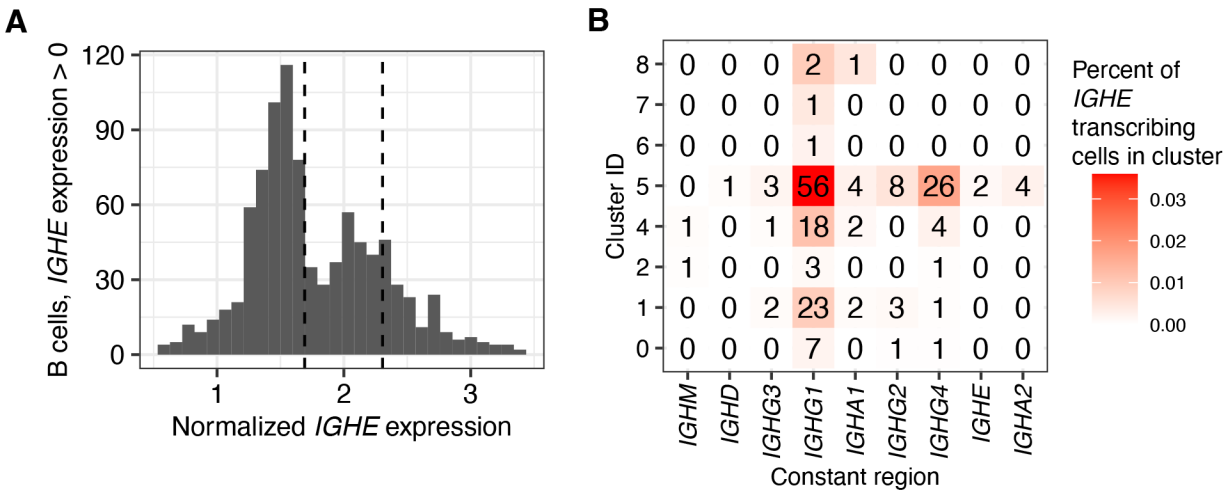

Fig. 3. *IGHE* transcribing B cells primarily use *IGHG1* and *IGHG4* BCR constant regions and are associated with cluster 5.

(**A**) Distribution of normalized *IGHE* expression among all B cells that express at least one *IGHE* transcript. Dashed lines show the amount of *IGHE* transcription in the two observed *IGHE*-constant region B cells. "*IGHE*-transcribing" cells are defined as cells with normalized *IGHE* expression at least as high as one of these two observed cells using *IGHE*-constant regions.(**B**) *IGHE*-transcribing cells were associated with each cluster and constant region. Cells of the grid are labeled by the total number of *IGHE*-transcribing cells and colored by the percentage of cells in each cluster that are *IGHE*-transcribing.

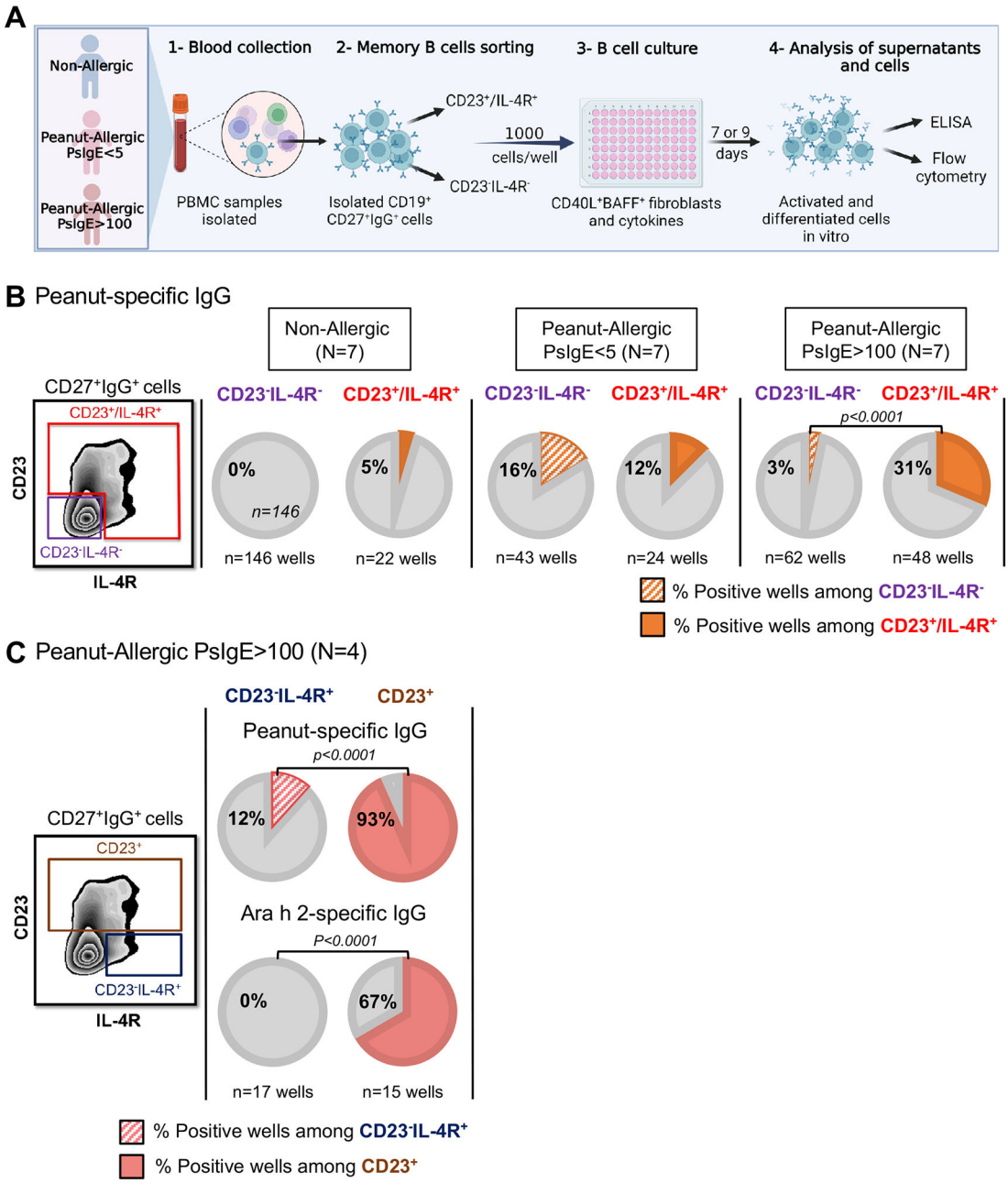

Fig. 4. IgG switched CD23$^+$ memory B cells from peanut-allergic individuals with PsIgE>100 showed an increased frequency of B cells producing peanut-specific IgG.

(**A**) Two distinct populations of CD27$^+$ memory B cells were sorted for memory B cell culture experiments. (**B**) Culture supernatants of IgG$^+$ memory B cells from non-allergic individuals (N=7, left), peanut-allergic individuals with PsIgE<5 (N=7, middle), and peanut-allergic individuals with PsIgE>100 (N=7, right) were tested for the reactivity against peanut by ELISA. Percent of positive wells among the total wells (indicated in each circle) of each culture is shown. Shaded orange pie slice represents a percent positive among CD23$^-$IL-4R$^-$ memory B cell culture wells and solid orange pie slice represents a percent positive among CD23$^+$/IL-4R$^+$ memory B cell culture wells. Differences in the percentage of positive wells

were determined by chi-square test. (**C**) Culture supernatants of IgG^+ memory B cells from peanut-allergic individuals with PsIgE>100 (N=4) were tested for the reactivity of IgG against peanut (top) and Ara h 2 (bottom) by ELISA. Percent of positive wells among the total wells (indicated in each circle) of each culture is shown. Shaded red pie slice represents a percent positive among $CD23^-IL\text{-}4R^+$ memory B cell culture wells and the solid red pie slice represents a percent positive among $CD23^+$ memory B cell culture wells. Differences in the percentage of positive wells were determined by chi-square test.

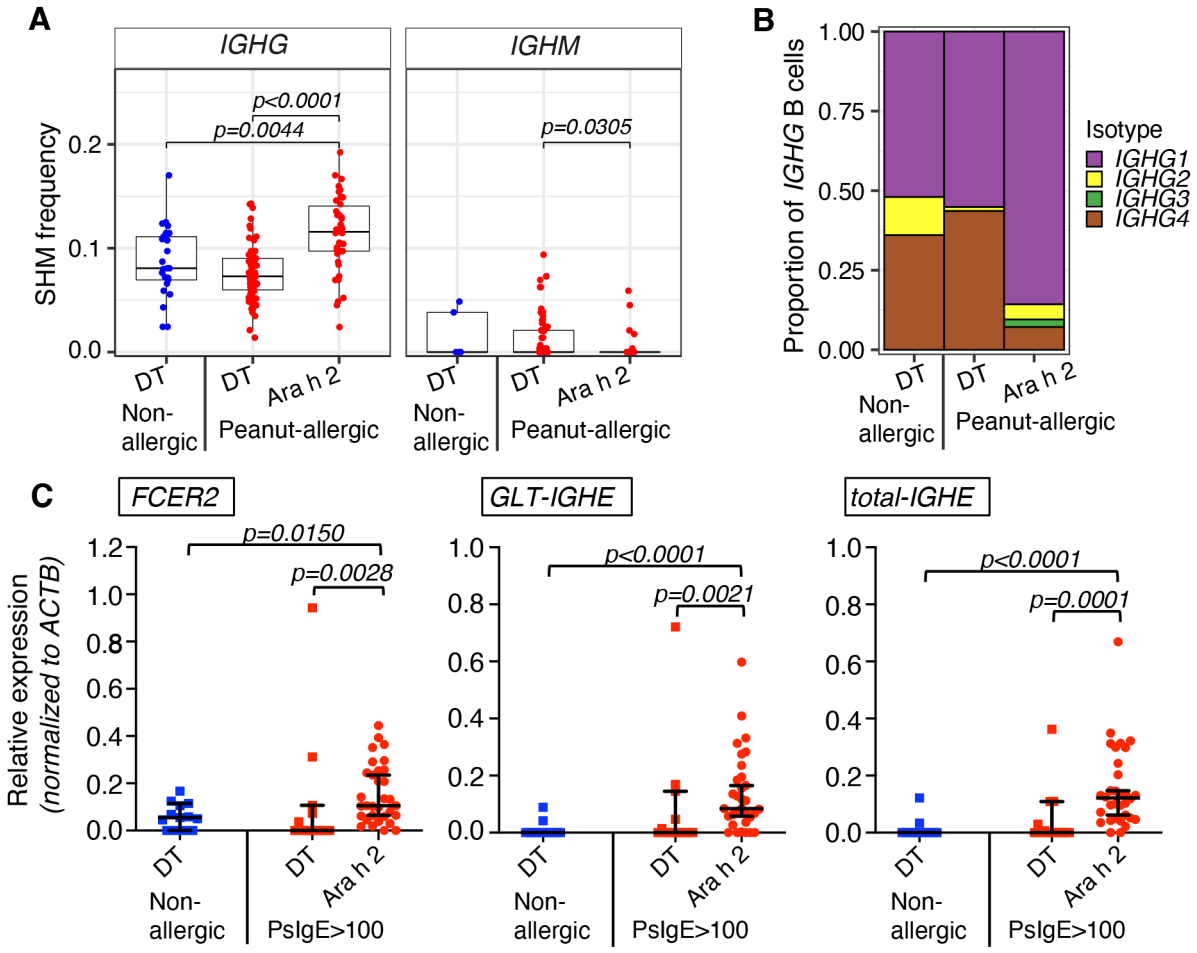

Fig. 5. BCRs from Ara h 2-binding IgG1$^+$ B cells are highly mutated, use *IGHG1* constant regions, and express *FCER2* and *IGHE*.

(**A**) Shown is the frequency of IGHV-gene somatic hypermutation for DT and Ara h 2-sorted B cells with either *IGHG* or *IGHM* constant regions from either peanut-allergic or non-allergic individuals. Data were analyzed by Mann-Whitney U test. Data are presented as box and whiskers plots in which boxes represent the median as well as first and third quartiles, and whiskers extend to the largest and smallest values within 1.5 times the interquartile range from the boxes. (**B**) Shown are the proportions of DT and Ara h 2-sorted *IGHG* B cells with associated *IGHG* sub-isotypes. (**C**) *FCER2, GLT-IGHE*, and total-*IGHE* gene expression was evaluated in single sorted IgG1$^+$ B cells. mRNA was quantified in single IgG1$^+$ B cells derived from Ara h 2 binders (n=34 cells) from 6 peanut-allergic individuals with PsIgE>100 and DT binders from 6 peanut-allergic individuals with PsIgE>100 (n=14 cells) and 3 non-allergic individuals (n=23 cells). Data are presented as mean±SD and were analyzed by Mann-Whitney U test.

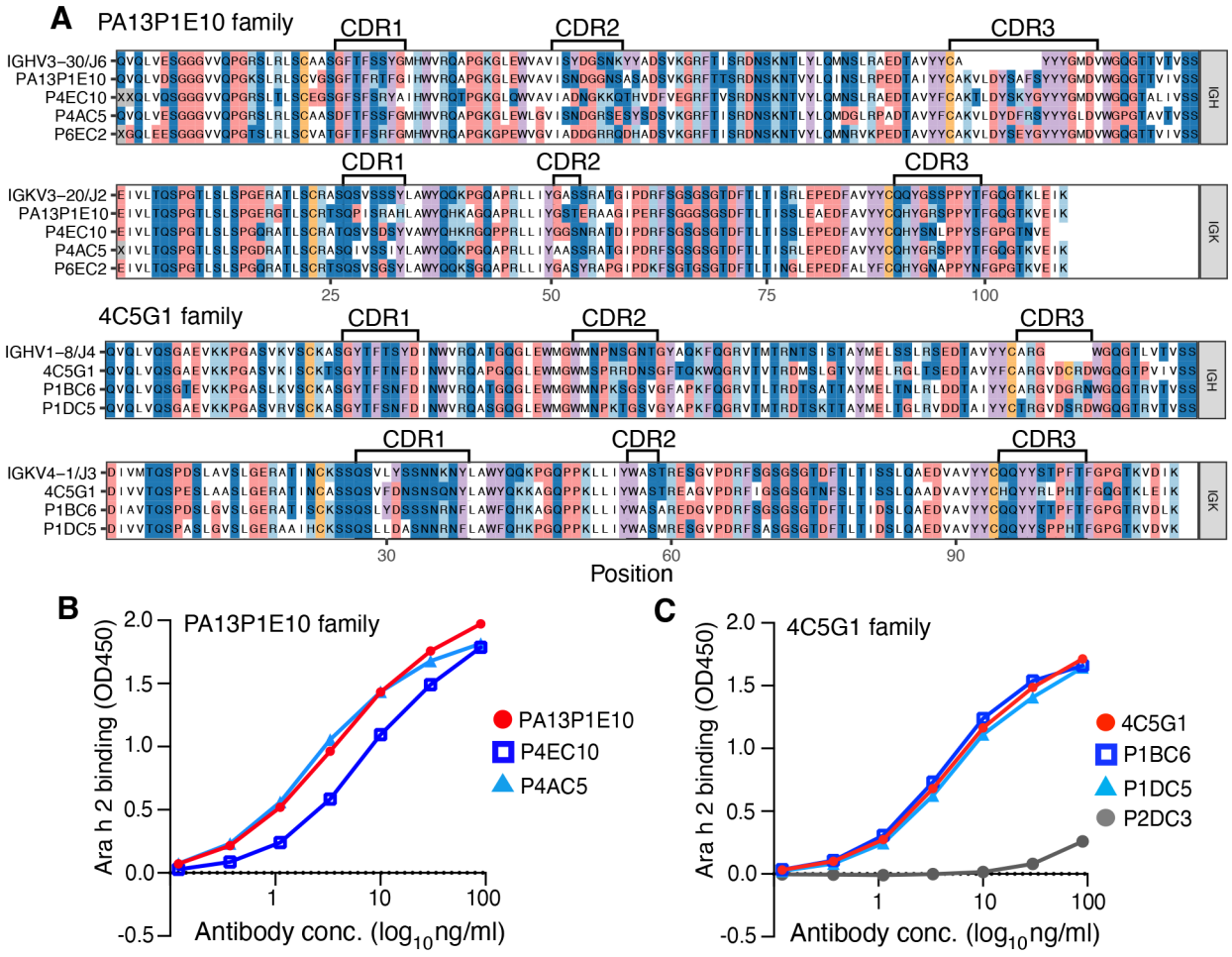

Fig. 6. Convergent Ara h 2-binding sequence clusters were observed across multiple donors.

(**A**) Amino acid alignment of heavy and light chain sequences is shown for two shared convergent sequence clusters. The upper sequence in each plot shows the predicted germline V and J gene sequences, followed by clonal members. Brackets indicate complementarity determining regions (CDRs) on sequence alignments. (**B and C**) ELISA curves are shown for monoclonal antibodies in the PA13P1E10 family (B) and the 4C5GF1 family (C) generated from V(D)J sequences of convergent clusters from (A).