

Allergen-induced activation of natural killer cells represents an early-life immune response in the development of allergic asthma

Matthew C. Altman, MD,^{a,b} Elizabeth Whalen, PhD,^a Alkis Togias, MD,^c George T. O'Connor, MD,^d Leonard B. Bacharier, MD,^e Gordon R. Bloomberg, MD,^e Meyer Kattan, MD,^f Robert A. Wood, MD,^g Scott Presnell, PhD,^a Petra LeBeau, ScD,^h Katy Jaffee, MS,^h Cynthia M. Visness, PhD,^h William W. Busse, MD,ⁱ and James E. Gern, MDⁱ
Seattle, Wash, Bethesda and Baltimore, Md, Boston, Mass, St Louis, Mo, New York, NY, Chapel Hill, NC, and Madison, Wis

Background: Childhood asthma in inner-city populations is a major public health burden, and understanding early-life immune mechanisms that promote asthma onset is key to disease prevention. Children with asthma demonstrate a high prevalence of aeroallergen sensitization and T_H2-type inflammation; however, the early-life immune events that lead to T_H2 skewing and disease development are unknown.

Objective: We sought to use RNA sequencing of PBMCs collected at age 2 years to determine networks of immune responses that occur in children with allergy and asthma.

Methods: In an inner-city birth cohort with high asthma risk, we compared gene expression using RNA sequencing in PBMCs collected at age 2 years between children with 2 or more aeroallergen sensitizations, including dust mite, cockroach, or both, by age 3 years and asthma by age 7 years (cases) and matched control subjects who did not have any aeroallergen sensitization or asthma by age 7 years.

Results: PBMCs from the cases showed higher levels of expression of natural killer (NK) cell-related genes. After cockroach or dust mite allergen but not tetanus antigen stimulation, PBMCs from the cases compared with the control subjects showed differential expression of 244 genes. This gene set included upregulation of a densely interconnected NK cell-like gene network reflecting a pattern of cell activation and induction of inflammatory signaling molecules, including the key T_H2-type cytokines *IL9*, *IL13*, and *CCL17*, as well as a dendritic cell-like gene network, including upregulation of CD1

lipid antigen presentation molecules. The NK cell-like response was reproducible in an independent group of children with later-onset allergic sensitization and asthma and was found to be specific to only those children with both aeroallergen sensitization and asthma.

Conclusion: These findings provide important mechanistic insight into an early-life immune pathway involved in T_H2 polarization, leading to the development of allergic asthma. (J Allergy Clin Immunol 2018;■■■:■■■-■■■.)

Key words: Asthma, allergens, transcriptomics, natural killer cells, dendritic cells

Asthma has emerged as a major public health problem in the United States over the past 25 years. Its incidence has increased over several decades, and it now affects more than 23 million Americans, including 7 million children, and costs US society \$56 billion annually.¹ The prevalence and severity of asthma are particularly high in low-income urban populations in the United States.^{2,3} Although asthma is a heterogeneous disease consisting of multiple endotypes, a substantial subset of asthma begins in childhood, and the immune changes leading to disease occur early in life.⁴⁻⁶ Currently, no prevention strategies exist, and therefore better understanding of the early-life immune changes that contribute to development of disease is critical.

From ^aBenaroya Research Institute Systems Immunology Division, Seattle; ^bthe Department of Medicine, University of Washington, Seattle; ^cthe National Institute of Allergy and Infectious Diseases, Bethesda; ^dBoston University School of Medicine; ^eWashington University School of Medicine, St Louis; ^fColumbia University College of Physicians and Surgeons, New York; ^gJohns Hopkins University School of Medicine, Baltimore; ^hRho, Chapel Hill; and ⁱthe University of Wisconsin School of Medicine and Public Health, Madison.

Supported in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, under contract nos. NO1-AI-25496, NO1-AI-25482, HHSN272200900052C, HHSN272201000052I, IUM1AI114271-01, and UM2AI117870. Additional support was provided by the National Center for Research Resources, National Institutes of Health, under grants RR00052, M01RR00533, IUL1RR025771, M01RR00071, IUL1RR024156, UL1TR000040, UL1TR001079, and 5UL1RR024992-02.

Disclosure of potential conflict of interest: G. T. O'Connor reports National Institutes of Health (NIH) institutional grant funding, has grants or grants pending from the NIH and Janssen Pharmaceuticals, and has received personal fees from AstraZeneca related to development of medications for asthma and COPD. L. B. Bacharier reports grants from the NIH/National Institute of Allergy and Infectious Diseases (NIAID) during the conduct of the study and received personal fees from Aerocrine, GlaxoSmithKline, Genentech/Novartis, Merck, Cephalon, DBV Technologies, Teva, Boehringer Ingelheim, AstraZeneca, WebMD/Medscape, Sanofi, Vectura, and Circassia outside the

submitted work. G. R. Bloomberg, P. LeBeau, and K. Jaffee report grants from the NIH/NIAID during the conduct of the study. M. Kattan reports grants from the NIH/NIAID during the conduct of the study and personal fees from Novartis Pharma outside the submitted work. R. A. Wood's institution has received NIH grant funding; has grants or grants pending from the NIH, DBV, Aimmune, Astellas, and HAL-Allergy; and reports unpaid consultancy to Stallergenes. C. M. Visness reports institutional NIH/NIAID grant funding. W. W. Busse reports grants from the NIH/NIAID during the conduct of the study and personal fees from Boston Scientific, ICON, Novartis, Glaxo SmithKline, Genentech, Roche, Boehringer Ingelheim, Sanofi Genzyme, AstraZeneca, Teva, 3M, PrEP Biopharm, Circassia, Regeneron, PeptinNovate, Knopp Bio, and Elsevier the submitted work. J. E. Gern reports grants from the NIH/NIAID during the conduct of the study and personal fees from Janssen, Regeneron, and PReP Biosciences and other funding from Boehringer Ingelheim outside the submitted work. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication October 18, 2017; revised January 29, 2018; accepted for publication February 12, 2018.

Corresponding author: Matthew C. Altman, MD, Systems Immunology Division, Benaroya Research Institute, 1201 Ninth Ave, Seattle, WA 98101. E-mail: maltman@benaroyaresearch.org.

0091-6749/\$36.00

© 2018 American Academy of Allergy, Asthma & Immunology

<https://doi.org/10.1016/j.jaci.2018.02.019>

Abbreviations used

DC:	Dendritic cell
DM:	Dust mite
FDR:	False discovery rate
iNKT:	Invariant natural killer T
IRIS:	Immune Response <i>In Silico</i>
KEGG:	Kyoto Encyclopedia of Genes and Genomes
NK:	Natural killer
TT:	Tetanus toxoid
URECA:	Urban Environment and Childhood Asthma

During childhood, allergic sensitization and associated T_H2 cell immune elements are key risk factors for airway inflammation, low lung function, airway hyperresponsiveness, and asthma.⁷⁻¹⁰ Developing sensitization to multiple allergens in the first few years of life is especially important in light of its poor prognosis with respect to persistent disease, increased exacerbations, and low lung function.¹¹⁻¹³ Although some genetic and environmental risk factors for T_H2 inflammation and asthma development are known,^{14,15} the early-life immune pathways that drive the imbalance toward a T_H2 immune response in human subjects are not well understood.⁶ Furthermore, not all subjects in whom allergen-specific T_H2 inflammation develops will experience asthma, and the systemic immunologic changes that specifically predispose to allergic asthma are unclear.

Several innate immune cells, including type 2 innate lymphoid cells, natural killer (NK) cells, invariant natural killer T cells (iNKTs), $\gamma\delta$ T cells, dendritic cells (DCs), monocytes, and others, can contribute to increased T_H2 -type cytokine signaling, and these cells are all key early responders to exogenous antigens and infections.¹⁶⁻¹⁹ Studies in mouse models have demonstrated the importance of innate immune cells in the development of allergic airway responses,²⁰⁻²⁴ and differences in type 2 innate lymphoid cell, NK cell, and iNKT cell populations have been observed in adults with established asthma.²⁵⁻²⁸ However, demonstration of the innate immune mechanisms that can initiate T_H2 asthma in human subjects is lacking.

To identify immunologic antecedents of allergic sensitization and asthma in early childhood, we studied children enrolled in an ongoing high-risk asthma birth cohort study: Urban Environment and Childhood Asthma (URECA).²⁹ Enrolled participants all have a parental history of allergic diseases or asthma and live in low-income census tracts. Per study design, the entire cohort had PBMCs collected at age 2 years and were stimulated with dust mite or cockroach extracts, tetanus toxoid (TT), or medium alone (nonstimulated). We used an unbiased whole-genome RNA sequencing approach coupled with cell deconvolution to investigate PBMC gene expression and stimulation responses, and here we report immunologic differences that can be detected at age 2 years and correspond to the development of aeroallergen sensitization and asthma.

METHODS**URECA population**

URECA is a birth cohort study initiated in 2005 in inner-city Baltimore, Boston, New York City, and St Louis, and details of the study design have been described elsewhere.²⁹ In brief, pregnant women age 18 years or older were recruited with selection criteria, including a history of asthma, allergic rhinitis, or eczema, in the mother or father. Between February 2005 and March 2007,

1850 families were screened, 776 met the eligibility criteria, and 560 newborns were enrolled at birth. Informed consent was obtained from the parent or legal guardian of the infant.

Maternal questionnaires were administered prenatally and postnatally every 3 months through age 7 years to ascertain wheezing illnesses and rhinitis symptoms. Allergen-specific IgE (ImmunoCAP; Phadia, Uppsala, Sweden) levels for milk, egg, peanut, and German cockroach were measured yearly until age 7 years. Specific IgE for DMs (*Dermatophagoides farinae* and *Dermatophagoides pteronyssinus*), dog, cat, mouse, and *Alternaria* species was added for ages 2 to 7 years. The lower limit of detection for the specific IgEs was 0.10 kU/L, and we manually set all values that were less than this to be 0.09 kU/L.

Skin prick tests were performed at 3, 5, and 7 years of age for 14 common aeroallergens. Allergen sensitization was defined as either a wheal of 3 mm or greater larger than that elicited by the saline control on skin prick testing or specific IgE levels of 0.35 kU/L or greater.

Asthma at age 7 years was defined by using a prespecified algorithm that included a parent-reported physician's asthma diagnosis, asthma symptoms, health care use for asthma, use of asthma medications in the previous year, spirometry with reversibility, and bronchial hyperresponsiveness assessed by means of methacholine inhalation challenge. Home visits 3 months after birth included house dust collection to assay for the allergens Bla g 1 (cockroach), Can f 1 (dog), Fel d 1 (cat), Der f 1 and Der p 1 (house DMs), and Mus m 1 (mouse) by means of ELISA (Indoor Biotechnologies, Charlottesville, Va).

PBMC processing and stimulation studies

Peripheral blood samples were collected at age 2 years and processed, as previously described.²⁹ In short, at each URECA research center, within 16 hours of blood collection, 1.0×10^6 mononuclear cells were separated and incubated (for 48 hours at 37°C in a 5% CO₂ atmosphere) in the presence of German cockroach extract (10 mg/mL; Greer Laboratories, Lenoir NC), DM extract (10 mg/mL; *Dermatophagoides pteronyssinus*; Greer Laboratories), TT (5 mg/mL; MassBiologics, Mattapan Mass), or medium alone (AIM-V media with human serum albumin; #12055-091; Invitrogen, Carlsbad, Calif). After incubation, cells were lysed with TRIzol (Tri-Reagent LS; Fisher Scientific, Waltham, Mass) and maintained at -80°C pending extraction of total cellular RNA.

Study design

An overview of the study design and analytic flow is presented in Fig E1 in this article's Online Repository at www.jacionline.org. For the initial case-control cohort (cohort 1), cases ($n = 21$) were URECA participants who met the following inclusion criteria: (1) 2 or more aeroallergen sensitizations at ages 3, 5, and 7 years, including sensitization to DM, cockroach, or both at each age; (2) diagnosis of asthma at age 7 years; and (3) PBMC RNA samples collected at age 2 years. Control subjects ($n = 30$) were URECA participants with (1) no aeroallergen sensitization through age 7 years, (2) did not have asthma at age 7 years, and (3) had PBMC RNA samples collected at age 2 years. Control subjects were selected to match to cases on at least a bivariate level, according to the following variables: sex, race, URECA site, birth season, and levels of Bla g 1, Der f 1, and Der p 1 measured in the home 3 months after birth. All available nonstimulated and cockroach-, DM-, and TT-stimulated PBMC samples from these 51 subjects were selected for RNA sequencing.

For cohort 2, 7 additional URECA participants were selected who met the following inclusion criteria: (1) 1 or more aeroallergen sensitizations by age 7 years, including cockroach sensitization; (2) a diagnosis of asthma at age 7 years; and (3) PBMC RNA samples collected at age 2 years. Fourteen additional control subjects were selected to match these cases by using the same criteria as the control subjects of cohort 1.

Furthermore, all URECA participants with PBMC RNA samples collected at age 2 years who had sufficient year 7 outcome data for group classification (full cohort) were added to the sequencing run to allow a 4-group comparison based on cockroach sensitization (age 3 years, yes/no) and asthma (age

7 years, yes/no). For the latter analyses, only nonstimulated and cockroach-stimulated PBMC samples were used for RNA sequencing.

RNA sequencing

RNA was extracted from TRIzol, according to the manufacturer's protocol. RNA quality was assessed by using RNA electrophoresis (Agilent, Santa Clara, Calif). All samples had RNA integrity number values of 7 or greater. Sequencing libraries were constructed from total RNA by using TruSeq RNA Sample Preparation Kits v2 (Illumina, San Diego, Calif) and clustered onto a flow cell by using a cBOT amplification system with a HiSeq SR v4 Cluster Kit (Illumina). Single-read sequencing was carried out on a HiSeq2500 sequencer (Illumina) by using a HiSeq SBS v4 Kit to generate 58-base reads, with a target of approximately 10 million reads per sample.

Samples were processed in 2 separate batches. Cohort 1 was processed as a single batch. All additional samples (cohort 2 and the full cohort) were processed as a second batch. One half of the samples from cohort 1 had sequencing library preparation and RNA sequencing repeated during the second batch to control for batch effects.

Bioinformatics analysis

FASTQ files were downloaded from <https://basespace.illumina.com>. Libraries were processed with Galaxy.³⁰ Libraries were aligned by using TopHat (version 1.4.1)³¹ to the human reference genome, Ensembl's *Homo sapiens* GRCh38, version 77 (GRCh38.77.gtf). The single-paired flag was set to "single," whereas all other TopHat parameters were set to defaults. The HTSeq-count³² was used to generate gene counts with mode as "intersection (nonempty)," and minimum alignment quality was set to 0 and otherwise set to default parameters.

Statistics

Genes with a trimmed mean of M values normalized³³ count of at least 1 in 20% of libraries and a gene biotype³⁴ of protein coding went into further analysis, leaving 13,867 Ensembl IDs that went into further analysis. Count data were converted into log₂ counts per million while weighting the observations to estimate the mean-variance relationship and incorporating the correlation between samples from the same subject by using the voom function³⁵ in limma.³⁶

Normalized voom counts were then used to test cases versus control subjects in nonstimulated samples by using linear modeling, including batch effects of site, sex, and year of blood draw. A linear model was built containing group (case vs control subject), stimulation (nonstimulation and cockroach stimulation), site, sex, year of blood draw, and home allergen levels (Bla g 1 in the bedroom and living room) and a random effect for the subject to test differences between cases and control subjects in response to cockroach stimulation. The equivalent was done for DM stimulation but with non-stimulation, DM stimulation, and Der f 1 and Der p 1 levels in the bedroom and living room. Where cell-specific gene sets were tested rather than individual genes, the geometric mean of all the genes within the set was used as a summary expression for the mixed-effect linear model. Multiple testing correction was done by using the Benjamini-Hochberg procedure, controlling the false discovery rate (FDR) at a level of 0.05.

Cell deconvolution to assess which cell types significantly affect gene expression was performed by using the Immune Response *In Silico* (IRIS) data set³⁷ as the reference data set to determine cell-specific marker genes.³⁸ Marker genes for B cells, DCs, monocytes, NK cells, CD4⁺ T cells, and CD8⁺ T cells were used. Surrogate cell proportions were calculated for our samples by using the digital sorting algorithm.³⁹ Finally, genes that were significant for cases versus control subjects in response to stimulation (cockroach or DM) were assigned to cell type by using the Subset Prediction from Enrichment Correlation algorithm.⁴⁰ The stability of the gene assignment to cell type was assessed through subsampling the marker genes through 100 iterations. Gene set correlation *P* values were calculated by simulating random gene sets of equivalent size over 10,000 simulations.

A general linear model was used with a gamma distribution and log link function to calculate a parameter estimate and *P* value to assess correlations among gene sets and clinical data. For comparisons of clinical and demographic data among groups, *P* values are calculated as appropriate for data: *t* test for normally distributed continuous data, Kruskal-Wallis test for nonnormally distributed continuous data, and either the χ^2 or Fisher exact test for categorical data.

Pathway analysis

Differentially expressed genes and gene sets were investigated by using Database for Annotation Visualization and Integrated Discovery, version 6.7.^{41,42} This database uses a modified Fisher exact test to identify specific biological/functional categories that are overrepresented in gene sets in comparison with a reference set (the default human genome was used as the reference set). Gene Ontology biological processes, canonical pathways from Kyoto Encyclopedia of Genes and Genomes (KEGG), and pathways from BioCarta that were associated with differentially expressed genes were used as databases.

STRING, version 10.0, which is a database of known and predicted protein-protein interactions, was used to determine interaction networks of cell-specific gene sets. Default parameters were used, including all interaction sources and using a minimum required interaction score of 0.40 to be included in the network.⁴³ Cytoscape, version 3.2.1, was used to draw interaction networks according to a prefuse force-directed layout by using the combined interaction score exported from STRING.⁴⁴ Unconnected nodes were excluded.

Data and material availability

RNA sequencing data were deposited in the National Center for Biotechnology Information Gene Expression Omnibus with accession number GSE96783.

RESULTS

Increased NK cell-type gene expression is associated with development of allergic asthma

Initially, we focused on 2 groups of children defined by clinical outcome in a nested case-control design (cohort 1). The case group was composed of all URECA participants who had 2 or more aeroallergen sensitizations, including DM, cockroach, or both, at ages 3, 5, and 7 years and had a diagnosis of asthma at age 7 years. We defined this group as having early-onset multiple sensitizations and asthma (Early group, *n* = 21). The control group (Neither group, *n* = 30) was composed of URECA participants who had no aeroallergen sensitization at age 3, 5, or 7 years and did not have asthma at age 7 years. These subjects were matched to the Early group according to demographic variables and allergen exposure in the home. RNA sequencing was performed on the year 2 nonstimulated, cockroach-stimulated, DM-stimulated, and TT-stimulated samples from these 51 subjects.

The Early group differed clinically from the Neither group, with a higher incidence of wheezing illnesses and higher Eczema Area and Severity Index scores in the first year of life, which are known risk factors for asthma development. At age 2 years, the time of blood draw, they had higher total IgE levels and higher specific IgE levels to food and airborne allergens, although most specific IgE levels were less than the cutoff for definition of sensitization (≤ 0.35 kU/L). By age 7 years, these same clinical differences were more pronounced, and the IgE levels markedly increased and demonstrated multiple specific sensitizations (see Table E1 in this article's Online Repository at www.jacionline.org).

Principal component analysis of gene expression of unstimulated PBMCs from these 2 groups showed similar global patterns among subjects (Fig 1, A). The primary source of variability in gene expression was city of residence (site); all 4 sites were represented equally in both groups. Assessment for differentially expressed genes comparing the Early and Neither groups while controlling for site demonstrated a small set of differentially expressed genes (13 genes at $FDR < 0.05$ and 131 genes at $FDR < 0.25$) between the groups, the majority of which (13/13 and 110/131, respectively) were expressed at a higher level in the Early group (Fig 1, B). The upregulated genes were highly enriched for NK cell–specific genes and no other cell type from the IRIS reference data set (Fig 1, C),³⁷ as well as the KEGG^{45,46} pathway NK cell–mediated cytotoxicity ($FDR = 2.7E-6$).

Gene expression of PBMCs from the same subjects after stimulation with the control antigen TT (to which all children had been vaccinated) demonstrated robust gene expression changes caused by stimulation (see Fig E2, A, in this article's Online Repository at www.jacionline.org) and no significant differences in gene expression between groups after stimulation (Fig 1, D), documenting a similar peripheral immune response to this common antigen to which children have uniform exposure. These results demonstrate that children with early-onset multiple sensitizations and asthma have greater gene expression of NK cell–related genes in resting PBMCs at age 2 years relative to matched control subjects but show similar responses to stimulation with a common vaccine antigen.

Allergen stimulation causes NK cell– and T_H2 -like gene activation

We next examined the cockroach- and DM-stimulated samples in cohort 1. Among all subjects, each allergen stimulation showed robust global gene expression relative to unstimulated samples (see Fig E2, B and C). The type of stimulation accounted for the greatest amount of variability among these samples, and each allergen stimulation showed a relatively distinct pattern of global gene expression (Fig 2, A).

Of the 21 Early group subjects, the 16 with cockroach sensitization by age 3 years (CR-Early group) demonstrated significant differences in expression of 206 genes in response to cockroach stimulation compared with the Neither group ($FDR < 0.05$; Fig 2, B). Similarly the 13 Early group subjects who were DM sensitized by age 3 years (DM-Early group) demonstrated significantly different expression of 91 genes in response to DM stimulation compared with the Neither group ($FDR < 0.05$; Fig 2, C). No genes were expressed differentially among the samples in association with total or allergen-specific IgE levels at age 2 years in linear models. Furthermore, a similar pattern of gene expression differences was observed between children in the Early group stratified by allergen sensitization at age 2 years, demonstrating these findings are independent of whether the Early group subjects were already sensitized at the time of blood draw.

The cockroach- and DM-induced gene lists showed enrichment for the same Gene Ontology biologic processes: immune response, cell proliferation, and programmed cell death. Genes in each of these pathways showed increased expression with both allergen stimulations in the Early group relative to the Neither group (Fig 2, B and C). These 2 gene lists, which are derived from different subsets of subjects and are induced by different allergen stimulations, show highly significant overlap ($P = 3.6E-75$, Fisher exact test;

Fig 2, D), suggesting the observed responses represent a common molecular response to allergen stimulation in susceptible subjects. Of all the 244 genes differentially expressed in the combined gene list, we focused on the 213 that showed increased expression in the Early group relative to the Neither group with both allergen stimulations (see Table E2 in this article's Online Repository at www.jacionline.org). None of these 213 genes were expressed differentially between groups after TT stimulation.

The set of allergen-upregulated genes specific to the Early group contains the same 13 genes that showed greater expression in the unstimulated samples (Fig 2, D); all these genes were upregulated further by allergen stimulation in the Early group (see Table E2). These 213 genes include numerous additional NK cell–related genes ($FDR = 6.9E-21$ for IRIS NK cell marker genes and $FDR = 2.8E-6$ for KEGG NK cell–mediated cytotoxicity), as well as several cytokine-signaling genes ($FDR = 7.5E-4$ for KEGG cytokine-cytokine receptor interaction), including multiple signaling genes implicated in T_H2 -type inflammation, such as *IL9*, *IL13*, *CCL17*, *CCL5*, and *CCR4*.

NK cells and DCs represent a primary source of response to allergen

The primary cellular source of the 244 differentially expressed gene transcripts was assessed by means of computational deconvolution with established markers for 6 major cell types contained in PBMCs (see Table E3 in this article's Online Repository at www.jacionline.org).^{40,47} One hundred eight of these 244 differentially expressed genes could be assigned to NK cells and were not associated with other cell types available for deconvolution (Fig 3, A, and see Fig E3 in this article's Online Repository at www.jacionline.org). All of these NK cell–assigned genes were increased in expression after cockroach or DM stimulation in the Early group compared with the Neither group. These NK cell–associated genes show a high level of gene-gene correlation ($r = 0.31$, $P < .0001$) and as a gene set show a coherent increase in expression level with both cockroach and DM stimulation in the Early group but not the Neither group (Fig 3, B).

These NK cell–associated genes represent a densely clustered network of known gene-gene interactions and associations (Fig 3, C) and are enriched for the KEGG pathway NK cell–mediated cytotoxicity ($FDR = 2.0E-5$), cytokine-cytokine receptor interaction ($FDR = 3.4E-5$), antigen processing and presentation ($FDR = 3.9E-5$), and the chemokine signaling pathway ($FDR = 5.3E-2$). The NK cell–type response includes both T_H1 - and T_H2 -type inflammatory genes but shows a particularly high increase in the critical T_H2 inflammatory cytokines *IL9*, *IL13*, and *CCL17* (Fig 3, D), suggesting that NK cell–derived T_H2 cytokines might be a principal source of skewing toward T_H2 inflammation at this developmental stage of the immune system. In contrast, expression of $CD4^+$ T cell–assigned genes was decreased predominantly by allergen stimulation (see Table E3) and was not correlated with levels of these T_H2 inflammatory cytokines.

A separate set of 52 of the 244 differentially expressed genes showed an expression pattern highly correlated with DC marker genes and not with other cell types (Fig 3, A, and see Fig E3). These DC-associated genes show a high level of gene-gene correlation ($r = 0.53$, $P < .0001$) and a moderate increase in response to cockroach allergen stimulation specific to the CR-Early group and not seen in the Neither group (see Fig E4, A, in this article's Online Repository at www.jacionline.org). In contrast, the DC

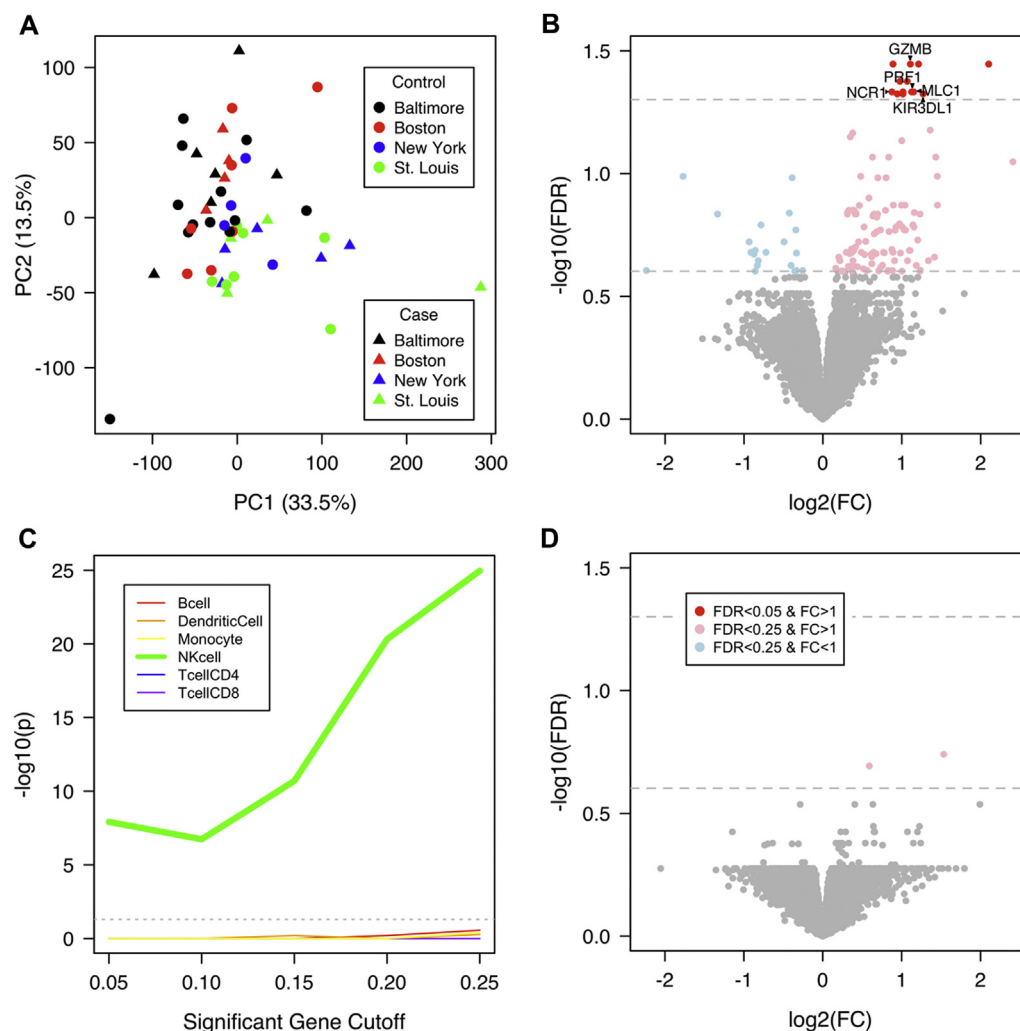


FIG 1. Differential expression of NK cell genes is associated with allergic asthma. **A**, Site but not case-control assignment accounted for the greatest source of gene expression variability in unstimulated PBMCs by using principal component (PC) analysis. **B**, Set of differentially expressed genes between the Early and Neither groups in unstimulated PBMC samples (nonstimulated linear model with fixed effects for site, sex, and year of blood draw; $n_{\text{Early}} = 19$, $n_{\text{Neither}} = 30$). Thirteen genes were significant at an FDR of less than 0.05 and fold change (FC) of greater than 1, as indicated by a red dot. Labeled genes (arrowheads) are NK cell marker genes in the IRIS database that have an FDR of less than 0.05. Genes with a semitransparent red dot have an FC of greater than 1 and an FDR of less than 0.25. Genes with a semitransparent blue dot have an FC of less than 1 and an FDR of less than 0.25. **C**, Genes that showed greater expression in the Early group compared with the Neither group were highly enriched for IRIS NK cell marker genes and no other cell type by using the Fisher exact test of overlap between significant genes and IRIS marker genes. **D**, There were no differentially expressed genes with FDRs of less than 0.05 between the Early and Neither groups after TT stimulation (tetanus stimulation linear model with fixed effects for site, sex, and year of blood draw; $n_{\text{Early}} = 19$, $n_{\text{Neither}} = 30$).

gene response to DM stimulation was not significant in the DM-Early subjects (see Fig E4, B). These DC-associated genes form a smaller network of known gene-gene interactions (see Fig E4, C) but do not show significant pathway-specific enrichment. The observed DC gene set includes the important lipid antigen-presenting genes *CD1B* and *CD1E*, both of which were induced highly by cockroach stimulation in the CR-Early group and also show an increase with DM stimulation in the DM-Early group (see Fig E4, D).

The remaining differentially expressed genes show correlations with B-cell, CD4⁺ T-cell, CD8⁺ T-cell, and monocyte marker genes, but they form smaller gene sets and do not show

enrichment for known molecular pathways. As gene sets, they show relatively modest differences between the Early and Neither groups with stimulation. The relevance of these differences is less evident and could either represent changes in cell composition, uncharacterized molecular pathways, or functional pathways that this study is underpowered to detect.

NK cell-type response is specific to allergy and asthma development

To assess the reproducibility of our findings and the specificity to the development of both allergy and asthma, we

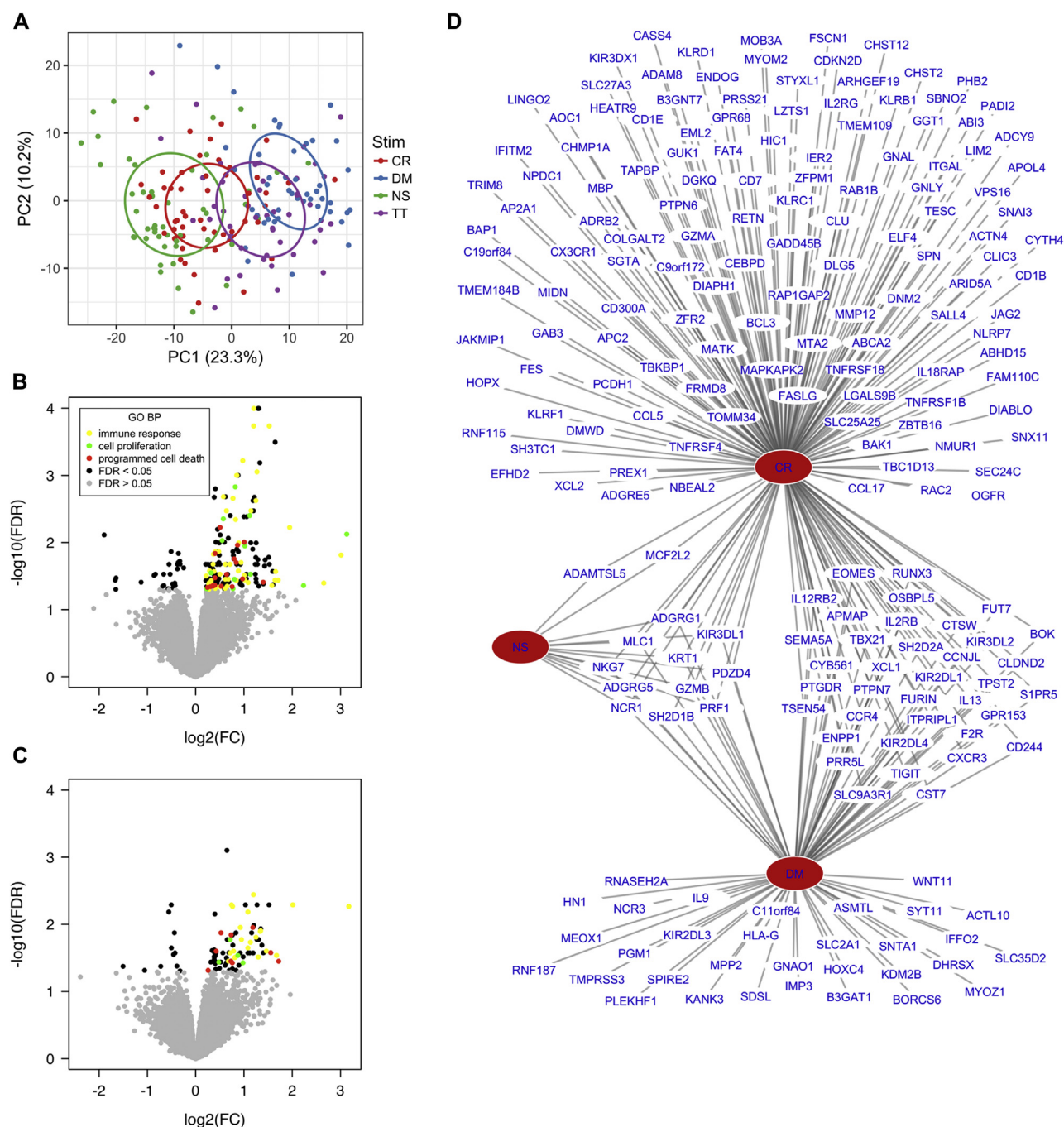


FIG 2. Allergen stimulation of PBMCs leads to robust gene expression changes. **A**, Stimulus accounts for the greatest source of gene expression variability in PBMCs by using principal component (PC) analysis. *El-lipses* represent 95% CIs. CR, Cockroach; NS, nonstimulated. **B**, CR-Early subjects demonstrated significant differences in expression of 206 genes in response to cockroach stimulation compared with the Neither group ($\text{FDR} < 0.05$). **C**, DM-Early subjects demonstrated significant differences in expression of 91 genes in response to DM stimulation compared with the Neither group ($\text{FDR} < 0.05$; stimulation linear modeling with fixed effects for site, sex, year of blood draw, home allergen levels, and stimulation and random effect for subject; $n_{\text{CR-Early}} = 16$, $n_{\text{DM-Early}} = 13$, $n_{\text{Neither}} = 30$). Differentially expressed genes in both comparisons are enriched for similar Gene Ontology biological process pathways. **D**, Significantly upregulated genes in the Early group in nonstimulated, cockroach-stimulated, and DM-stimulated samples show a high degree of overlap, as depicted in a bipartite network ($P = 3.6\text{E-}75$, Fisher exact test; cockroach- and DM-upregulated genes).

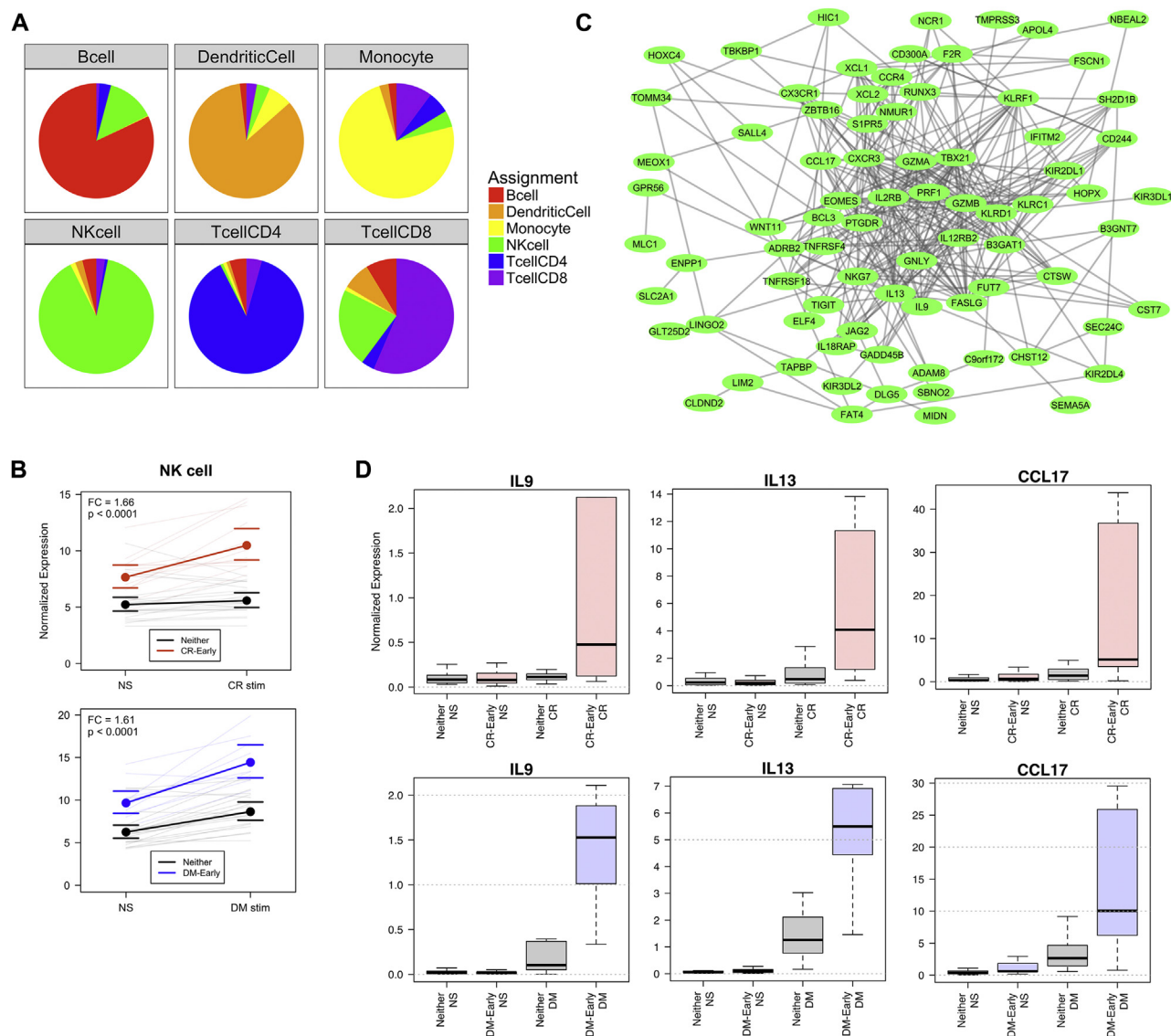


FIG 3. Allergen stimulation of PBMCs causes NK cell activation and T_H2 -like inflammation. **A**, The 244 genes could each be assigned to 1 of the 6 cell subsets by using marker gene cell deconvolution. Sensitivity analysis showed that gene assignments were relatively stable over 100 bootstrap simulations (Subset Prediction from Enrichment Correlation sensitivity analysis). **B**, The NK cell gene set shows significantly greater baseline expression in the Early group compared with the Neither group, which increases significantly with allergen stimulation in both the CR-Early and DM-Early groups; bounds represent 95% CIs (stimulation linear models with fixed effects for site, sex, year of blood draw, home allergen levels, and stimulation; random effect for individual subject; P value and fold change [FC] represent group term controlling for stimulation). **C**, NK cell genes represent a densely connected network of known gene-gene interactions. **D**, Box plot of normalized expression of *IL9*, *IL13*, and *CCL17* showing group differences with allergen stimulation (pink, CR-Early group; blue, DM-Early group; gray, Neither group). CR, Cockroach.

investigated the response to stimulation in an independently selected group of URECA participants (cohort 2). Cohort 1 focused on subjects with the most severe phenotype of interest, specifically those with multiple sensitizations by age 3 years and asthma by age 7 years. There were no additional URECA participants who met these selection criteria; however, there were participants with cockroach sensitization between ages 3 and 7 years and asthma by age 7 years. Seven such subjects were identified: 3 became cockroach sensitized by age 5 years, and the other 4 became cockroach sensitized by age 7 years; 6 of 7 had multiple allergic sensitizations at age 7 years. This group

was considered to have late allergy and asthma onset (Late group) and was compared with 14 newly selected and matched Neither group subject who did not experience allergic sensitization or asthma by age 7 years. Despite the later onset of allergen sensitization, a similar difference in gene expression was observed in PBMCs collected at age 2 years. In the Late group compared with the Neither group, unstimulated PBMCs and those stimulated with cockroach extract overexpressed the same NK cell gene set that was exhibited by the Early group (Fig 4, A and B). A difference in the DC gene set was not observed in this cohort.

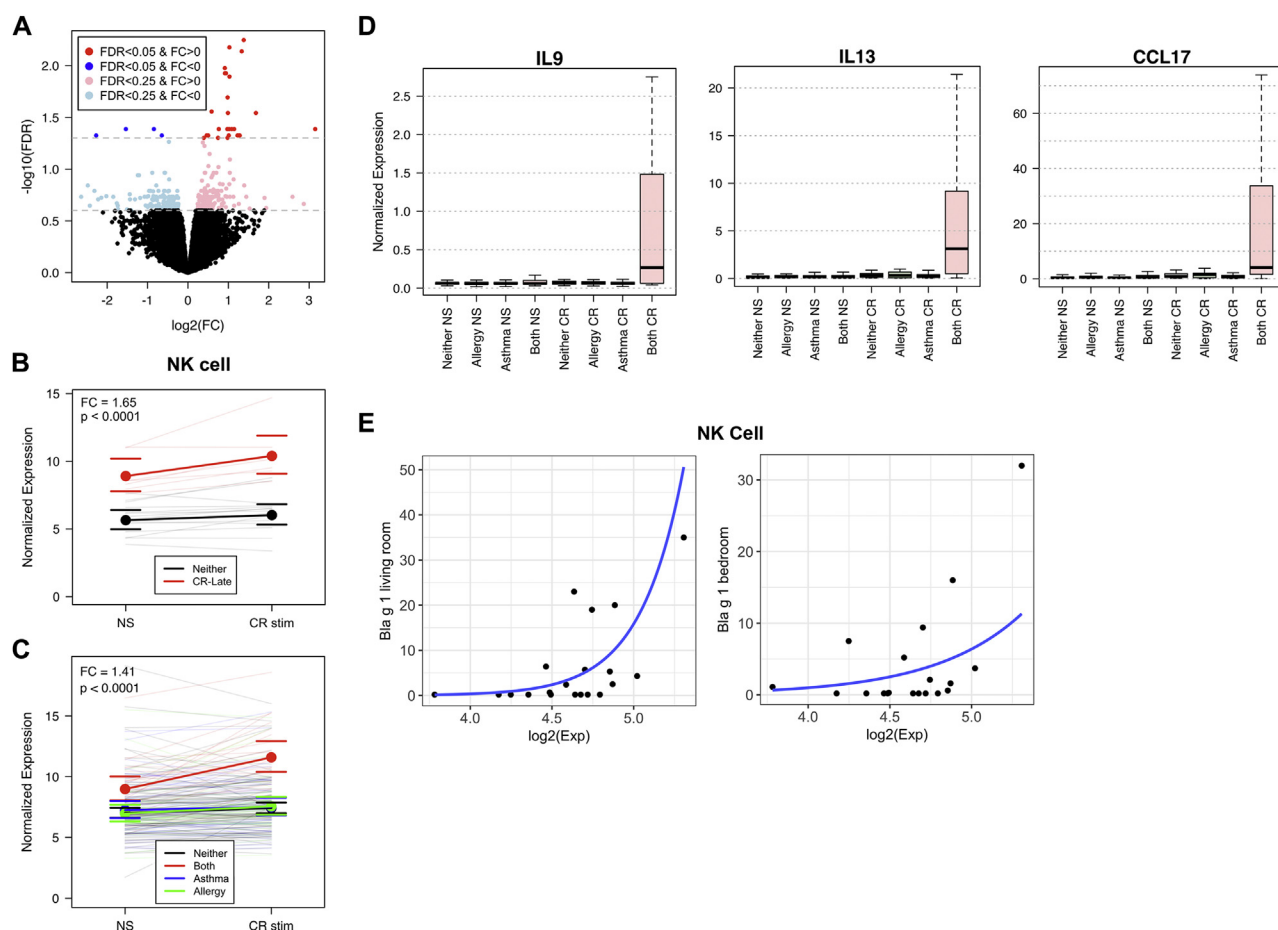


FIG 4. NK cell gene activation is specific to allergic asthma. **A**, The Late group demonstrated significant differences in expression of similar genes in response to cockroach (CR) stimulation compared with control subjects (31 genes at FDR < 0.05 and 326 genes at FDR < 0.25; stimulation linear model with fixed effects for site, sex, year of blood draw, home allergen levels, and stimulation and random effect for individual subject; $n_{\text{Late}} = 7$, $n_{\text{Neither}} = 14$). **B**, The NK cell gene set showed significantly greater baseline expression in the Late group compared with the Neither group, and expression increased significantly with cockroach stimulation in the Late group (P value and fold change [FC] represent group term controlling for stimulation; number of subjects in each comparison is represented in parentheses; bounds represent 95% CIs). **C**, The NK cell gene set shows significantly greater baseline expression in all cases (red; Both, Early and Late groups combined [$n = 20$]) compared with the Neither (black, $n = 102$), asthma-only (blue; Asthma, $n = 40$), and cockroach allergy-only (green; Allergy, $n = 35$) groups and increases with cockroach stimulation only in the Both group. **D**, Box plot of normalized expression of *IL9*, *IL13*, and *CCL17* showing group difference with allergen stimulation (pink, CR-Early; blue, DM-Early; gray, neither). **E**, Average expression value of the NK cell gene set showed correlation with Bla g 1 levels measured in the living room ($P < .05$) and bedroom ($P = .08$) in the Both group.

We next tested whether NK cell expression was related to allergy, asthma, or both. Children from the CR-Early and Late groups were combined with a single group ($n = 20$), and their nonstimulated and cockroach-stimulated gene expression data were compared with those of other subgroups of URECA participants classified by outcome at age 7 years: those with cockroach sensitization without asthma ($n = 35$), those with asthma without any allergic sensitization ($n = 40$), and those with neither allergic sensitization nor asthma ($n = 102$, see Table E4 in this article's Online Repository at www.jacionline.org). Increased NK cell gene expression at age 2 years was found only in children with both cockroach sensitization and asthma by age 7 years (Fig 4, C and D). Notably, similar results were found for increased expression of the DC gene set after

cockroach stimulation (see Fig E4, E). The other 3 groups were indistinguishable from one another in expression of NK and DC gene sets.

To better understand the cause of this NK cell gene set response among the Early and Late groups, we investigated the correlation of multiple clinical variables with NK cell gene set expression. The demographic variable that most highly correlated with expression level of the NK cell gene set was home cockroach allergen level measured in the first year of life. Among these 20 subjects, a higher level of cockroach exposure correlated with a higher level of NK cell gene set expression in response to cockroach stimulation (Fig 4, E). Home allergen levels did not correlate with NK cell gene expression within the other 3 groups. This finding suggests these disease-susceptible subjects might

have an NK response directly influenced by higher levels of allergen exposure early in life.

DISCUSSION

Children with multiple allergies in early life have an increased risk for wheezing and asthma. To identify immunologic correlates of early multiple allergic sensitization and asthma in high-risk urban children participating in the URECA study, we analyzed patterns of allergen-specific gene expression at age 2 years and compared these findings with longitudinal patterns of allergic sensitization and asthma assessed through age 7 years. Our results reveal differential gene expression in unstimulated and allergen-stimulated PBMCs collected at age 2 years in children with early multiple sensitization followed by asthma. Affected children exhibited increased expression of NK cell-related genes in unstimulated PBMCs, and differential expression of these NK cell-related genes was augmented by stimulation with cockroach and DM allergen but not TT. The largest set of differentially expressed genes after allergen stimulation was attributed to NK cells through cell marker deconvolution. This gene set describes a densely interconnected gene-gene interaction network that likely represents expansion of NK cells or a similar population of cytotoxic lymphocytes, as well as activation of these cells, leading to transcription of key immune signaling molecules, including the canonical T_H2 cytokines IL-9, IL-13, and CCL17. We also observed an upregulation of DC-related genes, including CD1 lipid presentation molecules, with cockroach stimulation.

Our results suggest that NK cell-like innate immune responses in early life contribute to development of aeroallergen sensitization-driven asthma. Mouse models have demonstrated a key role for NK cells in allergy and asthma development, and lack of NK cells attenuates airway inflammation in allergen-sensitized mice.^{23,24,48,49} Furthermore, children and adults with asthma have previously been observed to have higher levels of NK cells than nonasthmatic subjects after disease onset,^{28,50-52} and NK cells are a known source of T_H2 cytokines.^{53,54} Specifically, subpopulations of NK cells from adults with asthma have been shown to produce T_H2 cytokines⁵⁵ and also to directly regulate T-cell activation.⁵⁶ Our results are among the first evidence that NK cell-type responses in early life appear to play a direct role in the development of allergic asthma in children.

We have also found evidence of upregulation of lipid antigen presentation by DCs in response to cockroach, which is coincident with the observed NK cell response. Bidirectional NK-DC cross-talk can modify the differentiation of memory $CD4^+$ T-cell responses.⁵⁷⁻⁵⁹ In mouse models NK cells can also induce maturation of DCs and regulate their participation in allergen presentation and allergic sensitization.^{23,60-62} Our data implicate immune responses related to both NK cells and DCs in early-life immunopathogenesis, leading to allergic asthma.

Having observed the NK cell- and DC-type responses within this nested case-control cohort, we were able to demonstrate the same NK cell finding in a second group of children with allergic sensitization after age 3 years. Furthermore, we showed the specificity of these findings to children who have aeroallergen sensitization with asthma compared with those with only one of the conditions. Thus our results suggest that increased expression of NK-related genes occurs before allergic sensitization and predisposes specifically to asthma associated with allergic sensitization.

NK cells and DCs can differ intrinsically in children with allergic sensitization and asthma or might be modulated early in life by environmental factors. Given that these children tend to have more wheezing illnesses and viral respiratory tract infections in early life, as observed here and in previous work,⁶³ one possible mechanism is recruitment of NK cells from the periphery to the lung during those infections, as suggested by findings in adults with asthma,⁶⁴ where they could skew toward type 2 inflammation. However, viral infection alone seems insufficient to drive this NK signature because it was lacking in the group with asthma without allergic sensitization despite those subjects also having a higher incidence of early-life wheezing illnesses (see Table E4). We observed that among children with disease, the magnitude of the NK cell-type response correlated with cockroach allergen levels in house dust, suggesting allergen or associated exposures might play a key role in this immune response. Because cockroach allergen is also a marker of specific microbial exposures and the cockroach extract used in this study contains endotoxins and other bacterial components, it is possible that the microbial environment of these children also contributes to the observed immune response in light of known associations between the microbiome and disease development in this population.^{65,66} Whether NK cell responses are shaped by environmental, genetic, or epigenetic differences is under investigation. We have also not yet determined how NK cell responses change longitudinally during immune development and manifestation of clinical disease among these children. Assessing each of these factors will be critical to better understanding the mechanisms by which NK cells can contribute to disease development, improving identification of at-risk children early in life and ultimately identifying disease prevention strategies.

We used computational cell deconvolution to assess the source of the gene expression differences observed. A distinct advantage to assessing immune responses by using this method in a cell mixture is the ability to observe global expression patterns along with cell specificity and to infer potential cell-cell interactions.⁴⁷ However, although the gene sets described show high correlation with NK and DC cell marker genes, this assignment is probabilistic and limited by the ability to resolve cell types by patterns of marker gene expression in a complex mixture. Cells with highly similar transcriptional profiles, including NK cells, iNKT cells, and $\gamma\delta$ T cells, could each be contributing in part to the observed NK cell-type gene expression. In this study only bulk PBMC RNA had been stored, but in future studies, cell sorting followed by population or single-cell RNA sequencing would help resolve the source of the signal with greater specificity.

We conclude that children prone to aeroallergen sensitization and asthma have increased expression of NK cell-related genes in PBMCs in early life that is markedly upregulated on stimulation with allergen but not TT. This gene expression pattern includes transcription of *IL9*, *IL13*, *CCL17*, *CCL5*, and *CCR4*, which likely promote skewing toward T_H2 polarization. These findings are supported further by observed upregulation of DC genes after allergen stimulation, including those involved in lipid antigen presentation. These findings provide evidence for a mechanistic link for susceptibility to allergic asthma in high-risk children and suggest potential methods for early disease risk identification and disease prevention. Our findings also have the potential to identify key biomarkers for allergic asthma development, which will be necessary to conduct population-targeted asthma prevention studies.

We thank the entire URECA collaboration, which consists of the following institutions and investigators (principal investigators are indicated by asterisks): *Johns Hopkins University School of Medicine, Baltimore, Md*—R. Wood,* S. Leimenstoll, and R. Spellman; *Boston University School of Medicine, Boston, Mass*—G. O'Connor,* B. Caldwell, R. Cohen, E. Collins, S. Dorans, A. Florea, L. Gagalis, E. Gjerasi, D. Kretschman, M. Macharia, C. Sakow, S. Steinbach, A. Strakus, and B. West; *Columbia University Medical Center, New York, NY*—M. Kattan,* E. Arteago-Solis, A. Bello, Y. Fernandez-Pau, C. Lamm, S. Lovinsky-Desir, D. Perlaza, L. Peters, M. Pierce, C. Sanabia, S. Tsang, A. Valones, N. Whitney, and P. Yaniv; *Washington University School of Medicine and St Louis Children's Hospital, St Louis, Mo*—L. Bacharier,* A. Beigelman, G. Bloomberg, E. Tesson, and A. Freie; *National Institute of Allergy and Infectious Diseases, Bethesda, Md*—P. Gergen, A. Toggias, E. Smartt, and C. Czarniecki; *University of Wisconsin School of Medicine and Public Health, Madison, Wis*—J. Gern, W. Busse,* C. Sorkness, S. Doyle, R. Burton, P. Heinritz, K. Spring, A. Dresen, and S. Ramratnam; *Rho, Chapel Hill, NC*—S. Arbes,* C. Visness,* A. Calatroni, K. Jaffee, M. Yaeger, and H. Mitchell.

We thank all the families who have participated in the URECA study and D. Chaussabel and G. Nepom for their advice and discussions.

Key messages

- Inner-city children with allergy and asthma by age 7 years already show differences in PBMC gene expression at age 2 years, with (1) higher expression of cytotoxic lymphocyte genes in resting PBMCs and (2) upregulation of 2 distinct gene networks related to NK cells and DCs that include multiple T_H2-type cytokines.
- These findings are specific to the development of allergic asthma and not observed in children with only allergy or nonallergic asthma.

REFERENCES

- Barnett SBL, Nurmagambetov TA. Costs of asthma in the United States: 2002–2007. *J Allergy Clin Immunol* 2011;127:145–52.
- Weiss KB, Gergen PJ, Crain EF. Inner-city asthma. The epidemiology of an emerging US public health concern. *Chest* 1992;101:362S–7S.
- Crain EF, Weiss KB, Bijur PE, Hersh M, Westbrook L, Stein RE. An estimate of the prevalence of asthma and wheezing among inner-city children. *Pediatrics* 1994;94:356–62.
- Macaubas C, de Klerk NH, Holt BJ, Wee C, Kendall G, Firth M, et al. Association between antenatal cytokine production and the development of atopy and asthma at age 6 years. *Lancet* 2003;362:1192–7.
- Stern DA, Guerra S, Halonen M, Wright AL, Martinez FD. Low IFN- γ production in the first year of life as a predictor of wheeze during childhood. *J Allergy Clin Immunol* 2007;120:835–41.
- Bousquet J, Gern JE, Martinez FD, Anto JM, Johnson CC, Holt PG, et al. Birth cohorts in asthma and allergic diseases: report of a NIAID/NHLBI/MeDALL joint workshop. *J Allergy Clin Immunol* 2014;133:1535–46.
- Robinson DS, Hamid Q, Ying S, Tsicopoulos A, Barkans J, Bentley AM, et al. Predominant TH2-like bronchoalveolar T-lymphocyte population in atopic asthma. *N Engl J Med* 1992;326:298–304.
- Grunig G, Warnock M, Wakil AE, Venkayya R, Brombacher F, Rennick DM, et al. Requirement for IL-13 independently of IL-4 in experimental asthma. *Science* 1998;282:2261–3.
- Wills-Karp M. Interleukin-13: central mediator of allergic asthma. *Science* 1998;282:2258–61.
- Zhang D-H, Yang L, Cohn L, Parkyn L, Homer R, Ray P, et al. Inhibition of allergic inflammation in a murine model of asthma by expression of a dominant-negative mutant of GATA-3. *Immunity* 1999;11:473–82.
- Simpson A, Tan VY, Winn J, Svensen M, Bishop CM, Heckerman DE, et al. Beyond atopy: multiple patterns of sensitization in relation to asthma in a birth cohort study. *Am J Respir Crit Care Med* 2010;181:1200–6.
- Havstad S, Johnson CC, Kim H, Levin AM, Zoratti EM, Joseph CL, et al. Atopic phenotypes identified with latent class analyses at age 2 years. *J Allergy Clin Immunol* 2014;134:722–7.e2.
- Belgrave DC, Buchan I, Bishop C, Lowe L, Simpson A, Custovic A. Trajectories of lung function during childhood. *Am J Respir Crit Care Med* 2014;189:1101–9.
- Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. *Immunol Rev* 2011;242:10–30.
- Raedler LS, Schaub B. Immune mechanisms and development of childhood asthma. *Lancet Respir Med* 2014;2:647–56.
- Deniz G, van de Veen W, Akdis M. Natural killer cells in patients with allergic diseases. *J Allergy Clin Immunol* 2013;132:527–35.
- Walker JA, Barlow JL, McKenzie ANJ. Innate lymphoid cells—how did we miss them? *Nat Rev Immunol* 2013;13:75–87.
- Holt PG, Sly PD. Interaction between adaptive and innate immune pathways in the pathogenesis of atopic asthma. *Chest* 2011;139:1165–71.
- Subrata LS, Bizzintino J, Mamessier E, Bosco A, McKenna KL, Wikstrom ME, et al. Interactions between innate antiviral and atopic immunoinflammatory pathways precipitate and sustain asthma exacerbations in children. *J Immunol* 2009;183:2793–800.
- Lambrecht BN, Hammad H. Taking our breath away: dendritic cells in the pathogenesis of asthma. *Nat Rev Immunol* 2003;3:994–1003.
- Chang YJ, Kim HY, Albacker LA, Baumgarth N, McKenzie AN, Smith DE, et al. Innate lymphoid cells mediate influenza-induced airway hyper-reactivity independently of adaptive immunity. *Nat Immunol* 2011;12:631–8.
- Kim HY, Chang YJ, Subramanian S, Lee HH, Albacker LA, Matangkasombut P, et al. Innate lymphoid cells responding to IL-33 mediate airway hyperreactivity independently of adaptive immunity. *J Allergy Clin Immunol* 2012;129:216–27, e1–6.
- Mathias CB, Guernsey LA, Zammit D, Brammer C, Wu CA, Thrall RS, et al. Pro-inflammatory role of natural killer cells in the development of allergic airway disease. *Clin Exp Allergy* 2014;44:589–601.
- Ple C, Barrier M, Amniai L, Marquillies P, Bertout J, Tsicopoulos A, et al. Natural killer cells accumulate in lung-draining lymph nodes and regulate airway eosinophilia in a murine model of asthma. *Scand J Immunol* 2010;72:118–27.
- Akbari O, Faul JL, Hoyte EG, Berry GJ, Wahlström J, Kronenberg M, et al. CD4+ Invariant T-cell-receptor+ natural killer T cells in bronchial asthma. *N Engl J Med* 2006;354:1117–29.
- Shim J-U, Koh Y-I. Increased Th2-like invariant natural killer T cells in peripheral blood from patients with asthma. *Allergy Asthma Immunol Res* 2014;6:444.
- Barnig C, Cernadas M, Dutile S, Liu X, Perrella MA, Kazani S, et al. Lipoxin A4 regulates natural killer cell and type 2 innate lymphoid cell activation in asthma. *Sci Transl Med* 2013;5:174ra26.
- Duvall MG, Barnig C, Cernadas M, Rickelers I, Krishnamoorthy N, Grossman NL, et al. Natural killer cell-mediated inflammation resolution is disabled in severe asthma. *Sci Immunol* 2017;2.
- Gern JE, Visness CM, Gergen PJ, Wood RA, Bloomberg GR, O'Connor GT, et al. The Urban Environment and Childhood Asthma (URECA) birth cohort study: design, methods, and study population. *BMC Pulm Med* 2009;9:17.
- Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* 2016;44:W3–10.
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 2012;7:562–78.
- Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31:166–9.
- Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 2010;11:R25.
- Smedley D, Haider S, Durinck S, Pandini L, Provero P, Allen J, et al. The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Res* 2015;43:W589–98.
- Liu R, Holik AZ, Su S, Jansz N, Chen K, Leong HS, et al. Why weight? Modeling sample and observational level variability improves power in RNA-seq analyses. *Nucleic Acids Res* 2015;43:e97.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47.
- Abbas AR, Baldwin D, Ma Y, Ouyang W, Gurney A, Martin F, et al. Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* 2005;6:319–31.
- Chikina M, Zaslavsky E, Sealfon SC. CellCODE: a robust latent variable approach to differential expression analysis for heterogeneous cell populations. *Bioinformatics* 2015;31:1584–91.
- Zhong Y, Wan YW, Pang K, Chow LM, Liu Z. Digital sorting of complex tissues for cell type-specific gene expression profiles. *BMC Bioinformatics* 2013;14:89.
- Bolen CR, Uduman M, Kleinstein SH. Cell subset prediction for blood genomic studies. *BMC Bioinformatics* 2011;12:258.

41. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2008;4:44-57.
42. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2008;37:1-13.
43. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 2014;43:D447-52.
44. Shannon P. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13:2498-504.
45. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 2016;44:D457-62.
46. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;28:27-30.
47. Shen-Orr SS, Gaujoux R. Computational deconvolution: extracting cell type-specific information from heterogeneous samples. *Curr Opin Immunol* 2013;25:571-8.
48. Korsgren M, Persson CGA, Sundler F, Bjerke T, Hansson T, Chambers BJ, et al. Natural killer cells determine development of allergen-induced eosinophilic airway inflammation in mice. *J Exp Med* 1999;189:553-62.
49. Manners S, Alam R, Schwartz DA, Gorska MM. A mouse model links asthma susceptibility to prenatal exposure to diesel exhaust. *J Allergy Clin Immunol* 2014;134:63-72.
50. Krejsek J, Král B, Vokurková D, Derner V, Toušková M, Paráková Z, et al. Decreased peripheral blood gamma delta T cells in patients with bronchial asthma. *Allergy* 1998;53:73-7.
51. Lin SJ, Chang LY, Yan DC, Huang YJ, Lin TJ, Lin TY. Decreased intercellular adhesion molecule-1 (CD54) and L-selectin (CD62L) expression on peripheral blood natural killer cells in asthmatic children with acute exacerbation. *Allergy* 2003;58:67-71.
52. Timonen T, Stenius-Aarniala B. Natural killer cell activity in asthma. *Clin Exp Immunol* 1985;59:85-90.
53. Warren HS, Kinnear BF, Phillips JH, Lanier LL. Production of IL-5 by human NK cells and regulation of IL-5 secretion by IL-4, IL-10, and IL-12. *J Immunol* 1995;154:5144-52.
54. Loza MJ, Zamai L, Azzoni L, Rosati E, Perussia B. Expression of type 1 (interferon gamma) and type 2 (interleukin-13, interleukin-5) cytokines at distinct stages of natural killer cell differentiation from progenitor cells. *Blood* 2002;99:1273-81.
55. Wei H, Zhang J, Xiao W, Feng J, Sun R, Tian Z. Involvement of human natural killer cells in asthma pathogenesis: natural killer 2 cells in type 2 cytokine predominance. *J Allergy Clin Immunol* 2005;115:841-7.
56. Wingett D, Nielson CP. Divergence in NK cell and cyclic AMP regulation of T cell CD40L expression in asthmatic subjects. *J Leukoc Biol* 2003;74:531-41.
57. Jacobs B, Ullrich E. The interaction of NK cells and dendritic cells in the tumor environment: how to enforce NK Cell & DC action under immunosuppressive conditions? *Curr Med Chem* 2012;19:1771-9.
58. Van Elssen CHMJ, Vanderlocht J, Oth T, Senden-Gijsbers BLMG, Germeraad WTV, Bos GMJ. Inflammation restraining effects of prostaglandin E2 on natural killer-dendritic cell (NK-DC) interaction are imprinted during DC maturation. *Blood* 2011;118:2473-82.
59. Harizi H. Reciprocal crosstalk between dendritic cells and natural killer cells under the effects of PGE2 in immunity and immunopathology. *Cell Mol Immunol* 2013;10:213-21.
60. Deniz G, Akdis M, Aktas E, Blaser K, Akdis CA. Human NK1 and NK2 subsets determined by purification of IFN- γ -secreting and IFN- γ -nonsecreting NK cells. *Eur J Immunol* 2002;32:879.
61. Di Santo JP. Functionally distinct NK-cell subsets: developmental origins and biological implications. *Eur J Immunol* 2008;38:2948-51.
62. Hoshino T, Winkler-Pickett RT, Mason AT, Ortaldo JR, Young HA. IL-13 production by NK cells: IL-13-producing NK and T cells are present in vivo in the absence of IFN- γ . *J Immunol* 1999;162:51-9.
63. Jackson DJ, Gangnon RE, Evans MD, Roberg KA, Anderson EL, Pappas TE, et al. Wheezing rhinovirus illnesses in early life predict asthma development in high-risk children. *Am J Respir Crit Care Med* 2008;178:667-72.
64. Message SD, Laza-Stanca V, Mallia P, Parker HL, Zhu J, Keadze T, et al. Rhinovirus-induced lower respiratory illness is increased in asthma and related to virus load and Th1/2 cytokine and IL-10 production. *Proc Natl Acad Sci U S A* 2008;105:13562-7.
65. Lynch SV, Wood RA, Boushey H, Bacharier LB, Bloomberg GR, Kattan M, et al. Effects of early-life exposure to allergens and bacteria on recurrent wheeze and atopy in urban children. *J Allergy Clin Immunol* 2014;134:593-601.e12.
66. O'Connor GT, Lynch SV, Bloomberg GR, Kattan M, Wood RA, Gergen PJ, et al. Early-life home environment and risk of asthma among inner-city children. *J Allergy Clin Immunol* 2017 [Epub ahead of print].

Supplementary Materials

Fig. S1. Overview of study design and analysis flow. The schematic outlines the sample collection, sample selection and analyses conducted. URECA participants were enrolled prenatally. At age 2, PBMCs were collected and stimulated with allergen, antigen, and control medium. The current study was conducted after the children reached age 7. For the initial nested case-control cohort (cohort 1), RNA-sequencing was performed on the preserved year-2 PBMC RNA comparing “Early” cases (n=21) who developed ≥ 2 aeroallergen sensitizations by age 3 and a diagnosis of asthma at age 7 with “Neither” controls (n=30) who had no aeroallergen sensitization or asthma through age 7. Comparison was done by whole transcriptome differential gene expression followed by cell deconvolution of differentially expressed genes. To validate findings, cohort 2 included the remaining URECA participants who had ≥ 1 aeroallergen sensitization (including CR) by age 7 and a diagnosis of asthma at age 7, defined as “Late” (n=7), and compared to 14 additional “Neither” controls. A targeted analysis using the cell deconvolution assigned gene sets was performed. Finally all URECA participants with PBMC RNA samples collected at age 2 were compared in a 4 group comparison based on CR sensitization (age 3, yes/no) and asthma (age 7, yes/no). A targeted analysis using the cell deconvolution assigned was performed.

Fig. S2. Allergen and antigen stimulations of PBMCs lead to robust gene expression changes. (A) 7682 genes were differentially expressed comparing TT and NS samples in cohort 1. (B) 282 genes were differentially expressed comparing CR and NS samples in cohort 1. (C) 8009 genes were differentially expressed comparing DM and NS samples in cohort 1.(FDR<0.05

and $FC \geq 1.5$) (Stimulation linear modeling with fixed effects for site, gender, blood draw year, and stimulation and random effect for individual; $n=51$).

Fig. S3. Differentially expressed genes correlate with cell specific marker genes. The 244 genes assigned to one of the 6 cell subsets showed high positive correlation with marker genes specific to the respective cell type (Pearson correlations, red = positive correlation, blue = negative correlation, yellow = no correlation).

Fig. S4. Allergen stimulation of PBMCs cause increased DC gene expression. (A) The DC gene set increases significantly with CR stimulation in the CR-Early group compared to Neither. (B) The DC gene set is not significantly different among the CR-Early, DM-Early, or Neither groups with DM stimulation. Number of individuals in each comparison is represented in parentheses; bounds represent 95% confidence intervals (Stimulation linear models with fixed effects for site, gender, blood draw year, home allergen levels, and stimulation; random effect for individual). (C) DC genes represent a connected network of known gene-gene interactions. (D) Boxplot of normalized expression of *CD1B* and *CD1E* showing group differences with allergen stimulation (Pink=CR-Early, Blue=DM-Early, Grey=Neither). (E) The NK cell gene set shows significantly higher baseline expression in all cases (Red; “Both” = Early and Late groups, $n=20$) compared to Neither (Black, $n=102$), asthma only (Blue; “Asthma”, $n=40$), and CR allergy only (Green; “Allergy”, $n=35$) groups and increases with CR stimulation only in the Both group.

Table S1. Shown are the clinical and demographic characteristics of cohort 1.

Table S2. Shown are the 244 genes differentially expressed between cases and controls after CR and/or DM expression, including the fold change and FDR values and the cell type assignments.

Table S3. Shown are the cell specific marker genes derived from IRIS and used for SPEC cell type assignment.

Table S4. Shown are the clinical and demographic characteristics of the full cohort divided according to outcome at age 7. “Neither” are those who developed neither allergic sensitization nor asthma. “Both” are those from the CR-Early and Late groups combined into a single group, hence individuals who developed both CR sensitization and asthma. “Asthma only” are those who developed asthma without any allergic sensitization. “Allergy only” are those who developed CR sensitization without asthma.



URECA birth cohort – **Prenatal** enrollment
of children at risk for allergy/asthma



Year 2 unstimulated and
stimulated PBMCs collected



RNA-seq

Discovery Cohort 1
(nested case-control):

Year 7
Outcome



“Early” n=21
Year 3 multiple sensitizations
& Year 7 asthma

vs

“Neither” n=30
No sensitization or asthma at
any assessment through age 7



Differential gene expression
& Cell deconvolution

Validation Cohort 2:

“Late” n=7
Year 7 CR sensitization
& Year 7 asthma

vs

“Neither” n=14
No sensitization or asthma at
any assessment through age 7

Comparison of
Cell deconvolution gene sets

Full Cohort
4-group comparison:

“Both” n=20
Year 7 CR sensitization
& Year 7 asthma

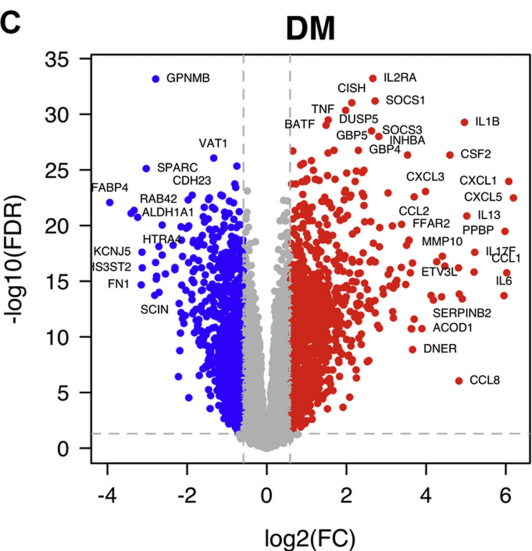
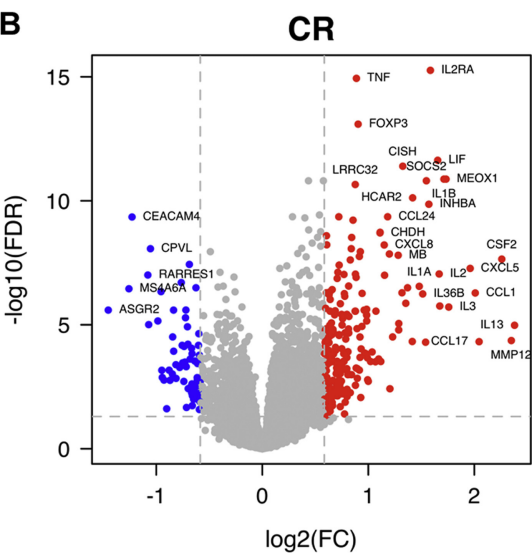
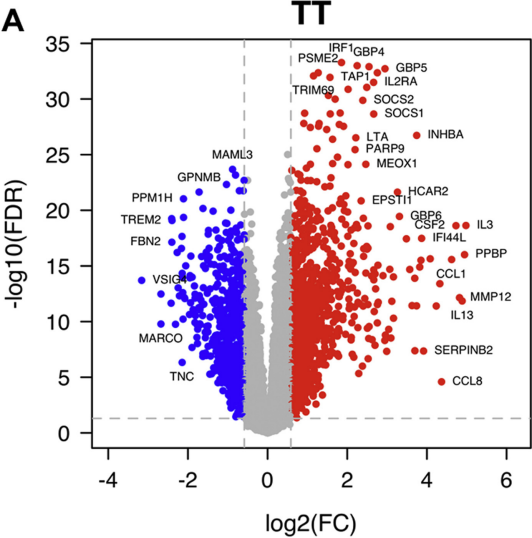
vs

“Neither” n=102
No sensitization or asthma at
any assessment through age 7

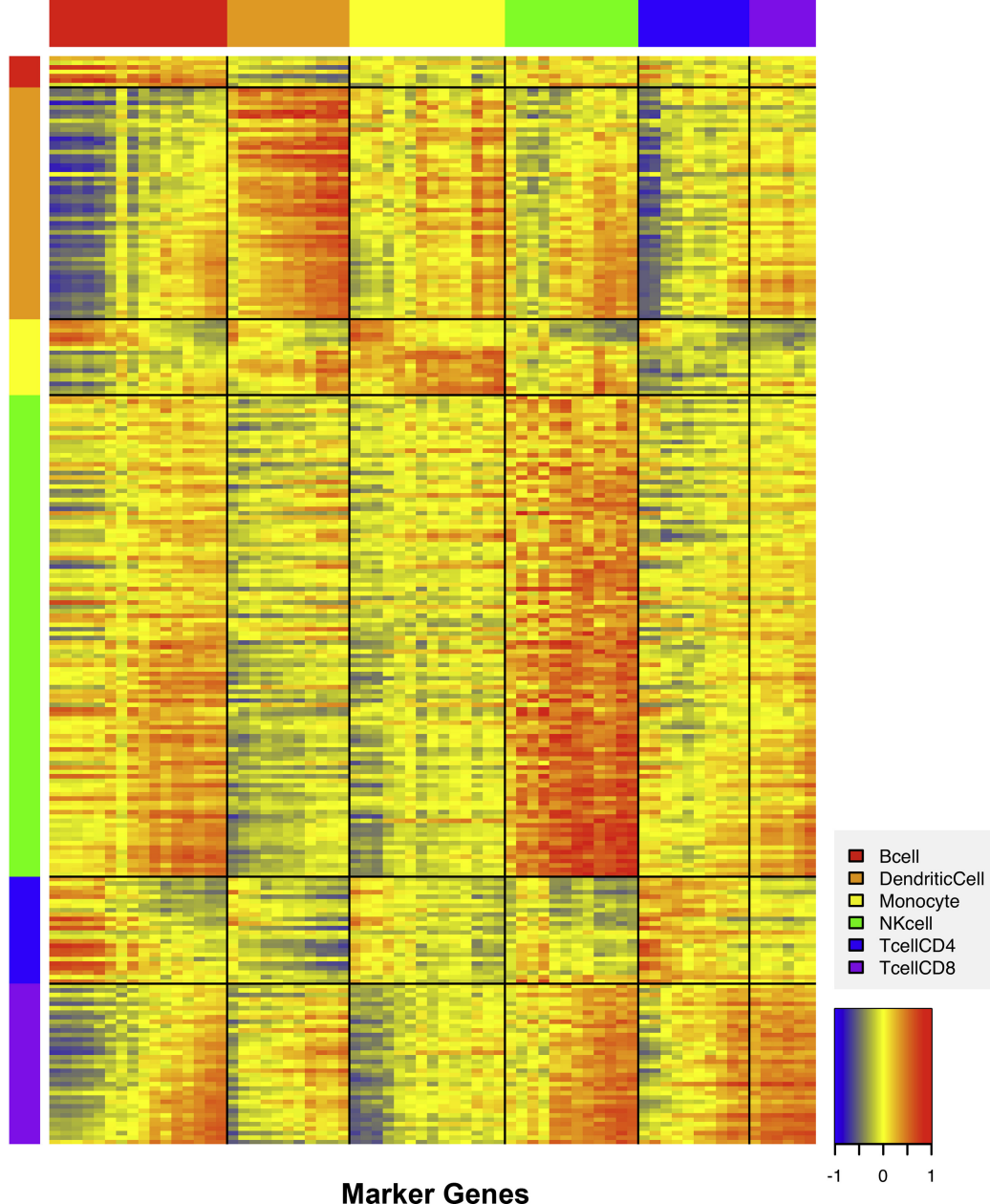
“Asthma only” n=40
No sensitization through age 7
Year 7 asthma

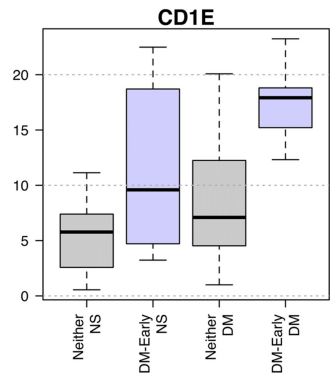
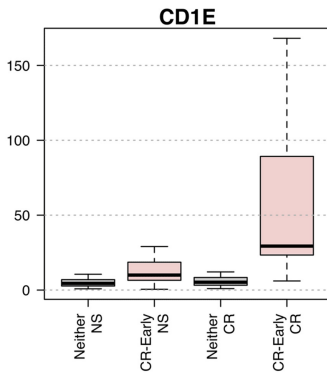
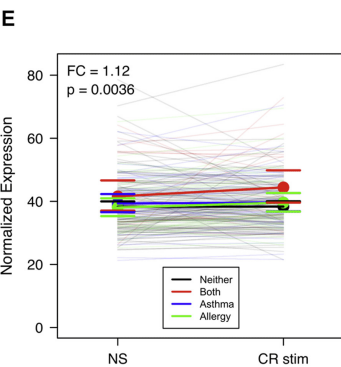
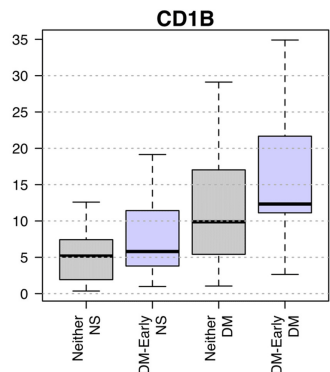
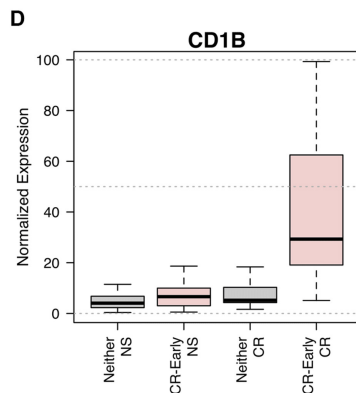
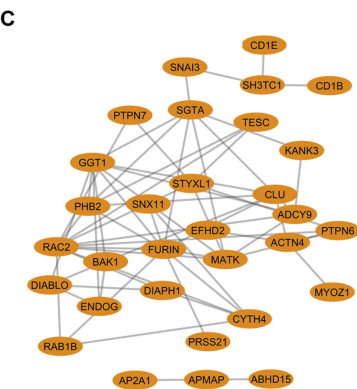
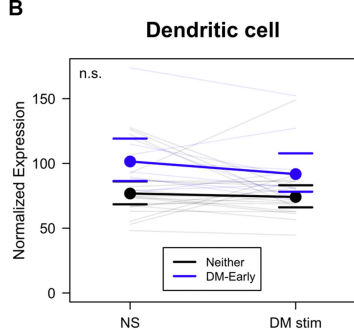
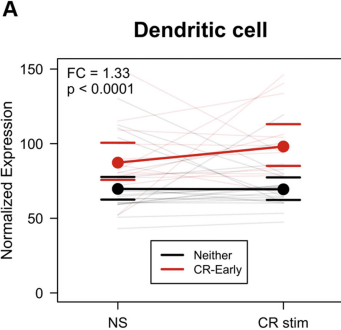
Comparison of
Cell deconvolution gene sets

“Allergy only” n=35
Year 3 CR sensitization
No year 7 asthma



244 differentially expressed genes





	n	Overall N=51	Controls N=30	Cases N=21	p- value†
Site: Baltimore	51	19 (37%)	12 (41%)	7 (33%)	0.82
Boston		11 (22%)	7 (23%)	4 (19%)	
New York		9 (18%)	4 (13%)	5 (24%)	
St. Louis		12 (23%)	7 (23%)	5 (24%)	
Mother's education: Less than high school	51	20 (39%)	15 (50%)	5 (24%)	0.18
High school		21 (41%)	10 (33%)	11 (52%)	
More than high school		10 (20%)	5 (17%)	5 (24%)	
Birth season: January – March	51	16 (31%)	9 (30%)	7 (33%)	0.81
April – June		8 (16%)	6 (20%)	2 (10%)	
July – September		15 (29%)	8 (27%)	7 (33%)	
October – December		12 (24%)	7 (23%)	5 (24%)	
Annual income < \$15K	51	34 (67%)	19 (63%)	15 (71%)	0.76
Mother is married	51	7 (14%)	3 (10%)	4 (19%)	0.43
Child's race: Black	51	42 (82%)	24 (80%)	18 (86%)	0.87
Hispanic		5 (10%)	3 (10%)	2 (9%)	
Mixed/other		4 (8%)	3 (10%)	1 (5%)	
Child's gestational age (wks.)	51	38.6 ± 1.7	38.6 ± 1.6	38.6 ± 1.7	0.92
Child's birth weight (g)	51	3197 ± 434	3145 ± 421	3271 ± 450	0.32
C-section delivery	51	15 (29%)	9 (30%)	6 (29%)	0.99
Mother breastfed at 3 months	48	9 (19%)	4 (14%)	5 (26%)	0.45
Age of mother at time of birth	51	23.1 ± 4.8	22.6 ± 4.8	23.8 ± 4.9	0.37
Child's gender: male	51	36 (71%)	18 (60%)	18 (86%)	0.09
Mother has had asthma	51	25 (49%)	13 (43%)	12 (57%)	0.49
Exposures in the first year of life					
Mother smoked	51	22 (43%)	15 (50%)	7 (33%)	0.37
Number of smokers in the home	51	1.2 ± 1.2	1.2 ± 1.4	1.1 ± 0.9	0.56
Der f – bedroom (ug/g)	49	0.28 [0.26; 0.35]	0.28 [0.26; 0.34]	0.28 [0.26; 0.45]	0.93
Der f – living room (ug/g)	41	0.27 [0.26; 0.29]	0.27 [0.26; 0.30]	0.27 [0.26; 0.29]	0.80
Der p – bedroom (ug/g)	49	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.33
Der p – living room (ug/g)	40	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.95
Bla g – bedroom (U/g)	47	0.59 [0.20; 12.0]	0.59 [0.20; 14.0]	0.24 [0.20; 4.4]	0.28
Bla g – living room (U/g)	41	0.20 [0.20; 5.7]	0.20 [0.20; 1.2]	1.7 [0.20; 6.2]	0.17
Clinical characteristics in first year of life					
Number of wheezing illnesses	51	0.7 ± 1.1	0.3 ± 0.7	1.2 ± 1.4	0.01
Any doctor visits for wheezing	51	6 (12%)	1 (3%)	5 (24%)	0.07
Any hospitalizations for wheezing	51	0 (0%)	0 (0%)	0 (0%)	n/a
Any wheezing	51	17 (33%)	5 (17%)	12 (57%)	0.01
Number of colds	51	4.4 ± 2.3	4.0 ± 2.3	4.9 ± 2.2	0.17
EASI score	51	1.8 ± 3.3	0.4 ± 0.7	3.7 ± 4.5	<0.01
Runny nose‡	50	18 (36%)	9 (30%)	9 (45%)	0.43
Stuffy nose‡	50	12 (24%)	6 (20%)	6 (30%)	0.51
Frequent sneezing‡	50	12 (24%)	4 (13%)	8 (40%)	0.04

Itchy eyes‡	50	10 (20%)	5 (17%)	5 (25%)	0.49
IgE levels at age 2 (kU/L)					
Total IgE	51	28.7 [10.2; 201]	16.5 [6.2; 31.3]	258 [130; 596]	<0.01
Peanut IgE	51	0.09 [0.09; 0.09]	0.09 [0.09; 0.09]	0.34 [0.09; 6.75]	<0.01
Milk IgE	51	0.09 [0.09; 0.21]	0.09 [0.09; 0.09]	0.23 [0.09; 1.57]	<0.01
Egg IgE	51	0.09 [0.09; 0.24]	0.09 [0.09; 0.09]	0.44 [0.16; 2.03]	<0.01
German Cockroach IgE	51	0.09 [0.09; 0.32]	0.09 [0.09; 0.09]	0.44 [0.09; 2.49]	<0.01
Dermatophagoides pteronyssinus IgE	51	0.09 [0.09; 0.17]	0.09 [0.09; 0.09]	0.19 [0.12; 1.33]	<0.01
Dermatophagoides farinae IgE	51	0.09 [0.09; 0.14]	0.09 [0.09; 0.09]	0.17 [0.09; 1.14]	<0.01
Cat IgE	51	0.09 [0.09; 0.09]	0.09 [0.09; 0.09]	0.16 [0.09; 0.64]	<0.01
Dog IgE	51	0.09 [0.09; 0.10]	0.09 [0.09; 0.09]	0.18 [0.09; 1.40]	<0.01
Mouse IgE	51	0.09 [0.09; 0.09]	0.09 [0.09; 0.09]	0.09 [0.09; 0.33]	<0.01
Alternaria IgE	51	0.09 [0.09; 0.09]	0.09 [0.09; 0.09]	0.09 [0.09; 0.17]	<0.01
IgE levels at age 7 (kU/L)					
Total IgE	47	77.0 [33.5; 508]	37.0 [16.0; 65.5]	564 [446; 1282]	<0.01
Peanut IgE	47	0.10 [0.09; 0.64]	0.09 [0.09; 0.10]	0.81 [0.15; 25.1]	<0.01
Milk IgE	47	0.12 [0.09; 0.23]	0.09 [0.09; 0.12]	0.26 [0.18; 1.12]	<0.01
Egg IgE	47	0.10 [0.09; 0.26]	0.09 [0.09; 0.10]	0.29 [0.12; 0.91]	<0.01
German Cockroach IgE	47	0.10 [0.09; 2.59]	0.09 [0.09; 0.10]	5.44 [1.69; 41.9]	<0.01
Dermatophagoides pteronyssinus IgE	47	0.10 [0.09; 2.47]	0.09 [0.09; 0.10]	3.41 [1.12; 8.36]	<0.01
Dermatophagoides farinae IgE	47	0.10 [0.09; 3.05]	0.09 [0.09; 0.10]	4.62 [0.75; 30.1]	<0.01
Cat IgE	46	0.10 [0.09; 0.92]	0.09 [0.09; 0.10]	4.00 [0.50; 20.2]	<0.01
Dog IgE	47	0.10 [0.09; 2.72]	0.09 [0.09; 0.10]	4.57 [0.99; 19.0]	<0.01
Mouse IgE	47	0.09 [0.09; 0.45]	0.09 [0.09; 0.10]	3.06 [0.09; 10.4]	<0.01
Alternaria IgE	47	0.10 [0.09; 0.14]	0.09 [0.09; 0.10]	0.24 [0.10; 2.78]	<0.01
Clinical characteristics at age 7					
Number of wheezing illnesses	51	0.9 ± 1.7	0.0 ± 0.0	2.2 ± 2.1	<0.01
Any doctor visits for wheezing	51	9 (18%)	0 (0%)	9 (43%)	<0.01
Any hospitalizations for wheezing	51	2 (4%)	0 (0%)	2 (10%)	0.16
Any wheezing	51	19 (37%)	0 (0%)	19 (91%)	<0.01
Number of colds	51	2.0 ± 2.4	1.6 ± 1.9	2.5 ± 2.9	0.21
EASI score	46	0.3 ± 1.1	0.0 ± 0.0	0.8 ± 1.7	0.06
Runny nose‡	50	8 (16%)	5 (17%)	3 (15%)	0.99
Stuffy nose‡	50	12 (24%)	6 (20%)	6 (30%)	0.51
Frequent sneezing‡	50	13 (26%)	6 (20%)	7 (35%)	0.39
Itchy eyes‡	50	9 (18%)	3 (10%)	6 (30%)	0.13
Any hay fever	50	10 (20%)	3 (10%)	7 (35%)	0.07

*Values are count (percentage), mean ± standard deviation, or median [IQR].

†P-values are from appropriate test for data – t-test for normally distributed, continuous data, Kruskal-Wallis test for non-normally distributed, continuous data, and either chi-squared or Fisher’s exact tests for categorical data.

‡Questions indicate routine occurrence, or greater than 2 days per week over the past year.

	n	Neither N=102	Both= CR Early+Late N=20	Asthma only N=40	Allergy only N=35	p- value†
Site: Baltimore	197	36 (35%)	7 (35%)	6 (15%)	12 (34%)	0.37
Boston		24 (24%)	4 (20%)	13 (33%)	8 (23%)	
New York		7 (7%)	4 (20%)	5 (13%)	3 (9%)	
St. Louis		35 (34%)	5 (25%)	16 (40%)	12 (34%)	
Mother's education: Less than high school	197	39 (38%)	4 (20%)	20 (50%)	15 (43%)	0.49
High school		41 (40%)	11 (55%)	13 (33%)	13 (37%)	
More than high school		22 (22%)	5 (25%)	7 (18%)	7 (20%)	
Birthseason: January – March	197	20 (20%)	5 (25%)	6 (15%)	7 (20%)	0.56
April – June		24 (24%)	2 (10%)	15 (38%)	8 (23%)	
July – September		28 (28%)	7 (35%)	11 (28%)	13 (37%)	
October – December		30 (29%)	6 (30%)	8 (20%)	7 (20%)	
Annual income < \$15K	197	63 (62%)	14 (70%)	29 (73%)	30 (86%)	0.06
Mother is married	197	15 (15%)	4 (20%)	3 (8%)	1 (3%)	0.12
Child's race: Black	197	78 (77%)	16 (80%)	33 (83%)	29 (83%)	0.82
Hispanic		15 (15%)	2 (10%)	6 (15%)	3 (9%)	
Mixed/other		9 (9%)	2 (10%)	1 (3%)	3 (9%)	
Child's gestational age (wks.)	197	38.7 ± 1.5	38.5 ± 1.9	38.8 ± 1.5	39.0 ± 1.5	0.67
Child's birth weight (g)	197	3226 ± 419	3314 ± 453	3187 ± 512	3260 ± 540	0.77
C-section delivery	197	27 (27%)	6 (30%)	12 (30%)	9 (26%)	0.96
Mother breastfed at 3 months	185	21 (22%)	5 (26%)	4 (11%)	10 (30%)	0.21
Age of mother at time of birth	197	24.7 ± 6.0	24.1 ± 5.1	24.2 ± 6.2	24.2 ± 5.9	0.95
Child's gender: male	197	49 (48%)	16 (80%)	24 (60%)	21 (60%)	0.05
Mother has had asthma	196	29 (28%)	10 (50%)	24 (62%)	16 (46%)	<0.01
Exposures in the first year of life						
Mother smoked	197	48 (47%)	6 (30%)	17 (43%)	11 (31%)	0.28
Number of smokers in the home	197	1.4 ± 1.3	1.2 ± 0.9	1.4 ± 1.3	1.1 ± 0.9	0.53
Der f – bedroom (ug/g)	196	0.27 [0.26; 0.33]	0.27 [0.26; 0.31]	0.27 [0.26; 0.28]	0.28 [0.27; 0.46]	0.05
Der f – living room (ug/g)	195	0.27 [0.26; 0.31]	0.27 [0.26; 0.29]	0.27 [0.26; 0.29]	0.28 [0.26; 0.36]	0.27
Der p – bedroom (ug/g)	195	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.46
Der p – living room (ug/g)	194	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.46 [0.46; 0.46]	0.59
Bla g – bedroom (U/g)	197	0.20 [0.20; 4.38]	0.85 [0.20; 5.78]	0.20 [0.20; 3.32]	3.60 [0.20; 17.0]	0.02
Bla g – living room (U/g)	197	0.20 [0.20; 3.13]	1.53 [0.20; 5.88]	0.98 [0.20; 7.35]	3.10 [0.20; 13.50]	0.04
Clinical characteristics in first year of life						
Number of wheezing illnesses	197	0.8 ± 1.9	1.2 ± 1.4	2.1 ± 2.4	1.5 ± 2.7	0.01
Any doctor visits for wheezing	197	6 (6%)	3 (15%)	13 (33%)	3 (9%)	<0.01
Any hospitalizations for wheezing	197	0 (0%)	0 (0%)	0 (0%)	0 (0%)	n/a
Any wheezing	197	37 (36%)	12 (60%)	29 (73%)	16 (46%)	<0.01

Number of colds	197	4.2 ± 2.9	5.3 ± 2.1	5.4 ± 2.6	4.3 ± 3.1	0.10
EASI score	197	0.7 ± 1.6	3.6 ± 4.5	1.0 ± 1.9	1.6 ± 3.2	<0.01
Sensitivity to food(s) – milk/egg/peanut	158	18 (21%)	8 (53%)	11 (34%)	11 (41%)	0.03
Clinical characteristics in third year of life						
Number of wheezing illnesses	197	0.4 ± 1.0	1.6 ± 2.4	2.3 ± 2.1	0.7 ± 1.4	<0.01
Any doctor visits for wheezing	197	5 (5%)	8 (40%)	11 (28%)	2 (6%)	<0.01
Any hospitalizations for wheezing	197	0 (0%)	1 (5%)	1 (3%)	0 (0%)	0.16
Any wheezing	197	25 (25%)	11 (55%)	32 (80%)	11 (31%)	<0.01
Number of colds	197	3.5 ± 3.0	4.0 ± 2.9	4.9 ± 3.1	3.2 ± 2.8	0.06
EASI score	180	0.6 ± 2.0	1.4 ± 2.0	0.3 ± 0.9	1.5 ± 6.9	0.40
Sensitivity to food(s) – milk/egg/peanut	175	26 (28%)	7 (44%)	17 (50%)	14 (42%)	0.10

*Values are count (percentage), mean ± standard deviation, or median [IQR].

†P-values are from appropriate test for data – t-test for normally distributed, continuous data, Kruskal-Wallis test for non-normally distributed, continuous data, and either chi-squared or Fisher’s exact tests for categorical data.