

Food allergy and omics



Gopal Krishna Dhondalay, PhD,^{a,d} Efren Rael, MD,^{a,d} Swati Acharya, PhD,^{a,d} Wenming Zhang, PhD,^{a,d}
Vanitha Sampath, PhD,^{a,d} Stephen J. Galli, MD,^{a,b,c} Robert Tibshirani, PhD,^{a,e} Scott D. Boyd, MD, PhD,^{a,b,d}
Holden Maecker, PhD,^{a,c,d} Kari Christine Nadeau, MD, PhD,^{a,d} and Sandra Andorf, PhD^{a,d} *Stanford, Calif*

Food allergy (FA) prevalence has been increasing over the last few decades and is now a global health concern. Current diagnostic methods for FA result in a high number of false-positive results, and the standard of care is either allergen avoidance or use of epinephrine on accidental exposure, although currently with no other approved treatments. The increasing prevalence of FA, lack of robust biomarkers, and inadequate treatments warrants further research into the mechanism underlying food allergies. Recent technological advances have made it possible to move beyond traditional biological techniques to more sophisticated high-throughput approaches. These technologies have created the burgeoning field of omics sciences, which permit a more systematic investigation of biological problems. Omics sciences, such as genomics, epigenomics, transcriptomics, proteomics, metabolomics, microbiomics, and exposomics, have enabled the construction of regulatory networks and biological pathway models. Parallel advances in bioinformatics and computational techniques have enabled the integration, analysis, and interpretation of these exponentially growing data sets and opens the possibility of personalized or precision medicine for FA. (*J Allergy Clin Immunol* 2018;141:20-9.)

Key words: Food allergy, bioinformatics, epigenomics, exposomics, genomics, transcriptomics, proteomics, metabolomics, microbiomics, integrative biology, systems biology

Food allergy (FA) is an abnormal, life-threatening immune response that occurs reproducibly on exposure to a certain food (www.foodallergy.org). These immune responses can be IgE dependent or non-IgE dependent and exclude celiac sprue, which has distinct immunologic mechanisms of response to gluten-containing foods.¹ The current standard of care for patients with FA is complete avoidance of the respective food or foods and the use of injectable epinephrine in cases of accidental exposure.²

Abbreviations used

AD: Atopic dermatitis
DNAm: DNA methylation
FA: Food allergy
GWAS: Genome-wide association study
MS: Mass spectrometry
OIT: Oral immunotherapy
TCR: T-cell receptor

The diagnosis and management of FA can also vary from one clinical practice setting to another.¹ Many immunotherapy trials have shown promise but are still experimental therapies and not readily available.³ FA directly affects patients and their families by decreasing quality of life in part because of anxieties associated with fear of adverse reactions on accidental ingestion of allergenic food and increased vulnerability of children to social ostracism.⁴

Reported cases of FA, as well as incidences of anaphylaxis, a potentially life-threatening condition, have been on the increase over the past few decades.⁵ In a recent retrospective cohort study of 333,200 children from the Children's Hospital of Philadelphia care network, FA prevalence was determined to be 6.7%.⁶ Based on the results of several studies, Sicherer and Sampson² estimated that food allergies likely affect nearly 5% of adults and 8% of children, and Andorf et al⁷ found that many patients with FA are allergic to multiple foods, with certain food allergies often occurring concomitantly. Although the reasons for the increasing prevalence of FA are unclear, it is thought to be influenced by multiple factors, such as genetic predisposition, increased hygiene, insufficient exposure to microbes, and changes in modern westernized diets.⁸ With the increasing prevalence of FA, there is an urgent need for better methods of diagnosis, prevention, and treatment.

From ^athe Sean N. Parker Center for Allergy and Asthma Research, ^bthe Department of Pathology, ^cthe Department of Microbiology & Immunology, ^dthe Institute for Immunity, Transplantation, and Infection, and ^ethe Department of Biomedical Data Sciences, Stanford University School of Medicine.

Authors are funded by AADCRC NIAID U19 grant and SNP Center for Allergy and Asthma Research at Stanford University.

Disclosure of potential conflict of interest: S. J. Galli has received personal fees from the National Institute of Allergy and Infectious Diseases Advisory Council; has received grants from the National Institutes of Health; and has a patent "Methods and kits for detecting basophil activation" (title of current draft of patent application) pending, and a patent "Multiplex isotype-specific antibody detection" (title of provisional patent application) pending. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication August 6, 2017; revised November 9, 2017; accepted for publication November 14, 2017.

Corresponding author: Kari Christine Nadeau, MD, PhD, Sean N. Parker Center for Allergy and Asthma Research at Stanford University, 269 Campus Dr, CCSR 3215, MC 5366, Stanford, CA 94305-5101. E-mail: knadeau@stanford.edu.

The CrossMark symbol notifies online readers when updates have been made to the article such as errata or minor corrections

0091-6749/\$36.00

© 2017 Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology

<https://doi.org/10.1016/j.jaci.2017.11.007>

TABLE I. Overview of omics approaches

Omics	Examples of method	Concept	Advantages	Discoveries with technology	Limitations
Epigenomics	Measurement of DNAm, other covalent DNA modifications, histone methylation, and other modifications	Assessment of nongenetic alterations of DNA or chromatin that affect gene expression and patient phenotypes	Might provide dynamic information about environmental influences on gene expression	Potential FA biomarkers and immunotherapy response biomarkers	Only one of many different mechanisms for gene regulation
Genomics	Genome/exome sequencing or microarray hybridization for detection of SNP and CNV variants for genotype-phenotype association studies	Association of genetic variants and inherited phenotypes; also somatic genetic rearrangements encoding antibodies and TCRs	Might help identify disease risk and disease prognosis and guide therapeutic selection; data are generally highly reproducible.	Discovery of genes involved in immune cell function; genetic associations with FA	Whole-genome methods have relatively high cost. Data capture a static feature of the patient. Associations with phenotype might be indirect.
Transcriptomics	RNA transcript detection and quantification by means of sequencing or microarray hybridization	Examination of expressed protein-coding or regulatory RNA molecules	Expressed genome is more closely associated with biological function and integrates genomic variants with environmental influences.	Genes expressed in immune system cells are associated with FA. Features of allergen biology are seen.	For cell type-specific data, sorting of subpopulations is required. Transcript levels might show variable correlation with protein levels.
Proteomics	MS detection and identification of peptides; potential peptide sequencing	Measurement of protein species and posttranslational modifications in cells or tissues of patients with FA	Expressed and modified proteins are a more direct measure of the biological state of a cell than genomic or transcriptomic measures.	Cellular and secreted proteins are involved in FA immune responses. Characterization of allergen proteins is seen.	Many proteins are more labile analytes than nucleic acids. Time course of protein changes can be affected by preformed stores versus new synthesis. This is a challenging methodology because of the diversity of analytes.
Metabolomics	MS or nuclear magnetic resonance spectroscopy of analytes	Provides a readout of metabolic pathway activity that can be associated with FA	Integrated analysis of host metabolic pathway activity and microbiome-derived metabolic products	Egg and peanut allergy metabolite pathways are seen. Urine metabolites in pregnancy and early life are associated with FA.	Some features of FA and therapeutic responses might be only indirectly associated with metabolic changes.
Microbiomics	Analysis of DNA, RNA, protein, and metabolites of microbes associated with host by using sequencing, proteomics, and metabolomics approaches	Analyze identity, frequency, and activity of microorganisms that can influence host biology	Microbes are likely to be key environmental mucosal factors modulating FA phenotypes.	Potential identification of disease risk factors and biomarkers for disease	Interactions between microbiome and host cells can occur often in less-accessible tissue sites, such as the gut.
Exposomics	Measurement of environmental exposures using wearable technology, sensors, and surveys	Accurate measurement of environmental exposures of an organism is critical for understanding environmentally influenced diseases, such as FA.	Some estimate that genetic influence on FA account for only 30% of disease risk. Interactions with the environment play a major role in disease manifestation. ^{9,10}	Risk factors contributing to FA incidence and prevalence	Standardization of these diverse measurements is still a work in progress.

CNV, Copy number variation; SNP, single nucleotide polymorphism.

WHAT ARE OMICS?

Omics refers to systematic and comprehensive fields of study in which advanced and generally high-throughput technologies are used to investigate, characterize, and quantify the molecules and

cells that comprise an organism and to understand their interactions and relationships.

Specifically, studies of proteins, small molecules, transcripts, genes, microbes, and environmental exposures have created the

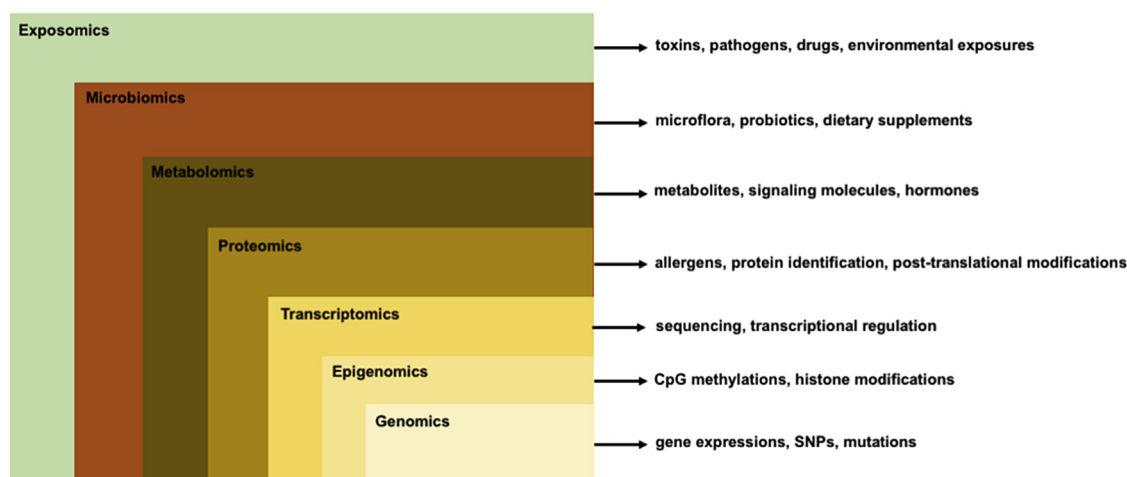


FIG 1. Omics technologies. Representation of the different omics technologies and their scope of application. *SNPs*, Single nucleotide polymorphisms.

burgeoning group of omics sciences termed proteomics, metabolomics, transcriptomics, genomics, microbiomics, and exposomics, respectively. Table I^{9,10} and Fig 1 provide information about some common omics technologies.

With the advent of advanced high-throughput technologies, it is now possible to simultaneously characterize and quantify a large number of analytes or cellular components from a single biological sample.¹¹ In contrast to traditional biological studies, which emphasize hypothesis-driven and reductionist experimentation, omics approaches use technologies that create large data sets that can be empirically analyzed by using unbiased approaches (ie, independent of specific hypotheses) to reveal findings that might generate hypotheses.¹² With further refinements of these omics technologies, the bottleneck impeding new discoveries is no longer the laborious generation and collection of biological and medical data but rather the ability to identify and interpret the underlying information in these large and often noisy data sets.¹³ By integrating data from different omics technologies (termed systems biology), one can obtain a high-dimensional view of biological processes to understand the layered complexity of and interactions among the different biological systems that influence disease processes.¹⁴

Omics technologies are now used commonly in diverse areas of research because of their increasing availability and decreased costs (Table II).¹³ There are many examples of diagnostic biomarkers and therapeutic strategies suggested by omics research and systems biology approaches.¹⁵ For example, in patients with breast cancer, high-throughput gene expression profiling studies have enabled the development of commercial prognostic kits, such as MammaPrint-DX, Oncotype-DX, EndoPredict, and Prosigna, to aid clinical decision.¹⁶

In patients with FA, successful oral immunotherapy (OIT) in patients with peanut allergy has been associated with conversion of allergen-specific T cells to anergic T cells.¹⁷ These findings, along with technological developments that allow real-time monitoring, could have therapeutic implications and enable personalized medicine.¹⁷ Fig 2 schematically represents the general workflow of data collection, analysis, and interpretation that can be used to generate information that might be useful in guiding clinical decisions for personalized medicine. Table III lists

what is known and unknown of omics in FA. Although the use of omics technologies and approaches in FA has thus far been limited, its potential is being recognized.¹⁸

In this review we provide a current and broad overview of some of the commonly used omics technologies and corresponding data analysis approaches. We focus specifically on studies that have used omics in FA research and discuss the potential of such approaches in future research.

OMICS TECHNOLOGIES IN FA

Protein level

Proteins, functional components of cells that catalyze vital processes, and their characterization specific to cells, tissues, or organisms is termed the proteome. Comprehensive investigation of these proteins is called *proteomics*, which encompasses the study of a protein's structure, cellular interactions, functions, localization, and posttranslational modifications. Using proteomic approaches to identify, characterize, and quantitate food allergens is now referred to as *allergenomics*,¹⁹ and there are many databases with information on the nomenclature and structures of food and environmental allergenic proteins (<http://www.allergenonline.org/>, <http://www.allergen.org/>, <http://fermi.utmb.edu/>, and <http://www.allergome.org/>).

With the aid of proteomics, the primary structure of more than 850 allergens has been identified,²⁰ including food allergens such as those found in rice endosperms,²¹ legumes (peanuts, soybeans, and lentil), and sesame.²² Proteomics has also enabled better understanding of cross-reactivities among allergens. By using matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MS), 4 peptides from α -casein (cow's milk allergen) and 3 peptides from Gly m 5 (soy milk allergen) were identified as having a common core motif, suggesting that these epitopes are likely responsible for the cross-allergenicity observed between cow's milk and soy milk.²³ Furthermore, bovine α -S1-casein was identified in human colostrum as a possible cause of allergic sensitization to cow's milk in infants who were exclusively breast-fed.²⁴

Proteomics can be used to interrogate the host response to FA. Traditional approaches use peptide-based MS to discover, in a

TABLE II. Estimated costs associated with omics technologies

Omics	Methods	Overview of workflow	Estimated cost of experimental procedures (20 samples)	Total timeline (omics service providers)
Epigenomics	A. Genome-wide methylation analysis	1. FFPE DNA extraction	\$1,500 (DNA extraction + bisulfite conversion)	3 mo
	B. Methylation array (27,000/450,000/850,000 CpG sites)	2. Bisulphite conversion	\$30,000 (sequencing + genome-wide analysis)	
	C. Targeted sequencing for DNAm analysis	3. Pyrosequencing	\$10,000-\$20,000 (sequencing + array-based analysis)	
	● Primer design and validation	4. High-throughput sequencing	\$10,000-\$20,000 (sequencing + targeted analysis)	3 mo
	● Targeted amplification	5. DNAm analysis		
	● Adapter ligation and barcoding			
	● Sequencing (50-3,000 loci)			
	D. RNA methylation (similar method as DNA)			3 mo
	E. ChIP-seq	1. Immunoprecipitated DNA sequencing	Type of IP antibody-dependent variable costs	
		2. Library preparation	\$2,000 (library)	
		3. Peak calling and chromatin analysis	\$4,000 (sequencing 20 pooled samples)	
	F. ATAC-Seq	1. Cell harvest and lysis		
		2. Crude nuclei harvest, Tn5 tagmentation	\$2,000 (tagmentation, DNA tagging, purification)	3 mo
		3. DNA tagging and purification	\$4,000 (library and agarose gel-based purification)	
		4. Amplification and library purification	\$4,000 (sequencing 20 pooled samples)	
Genomics	A. DNA sequencing for SNPs and mutations	5. High-throughput sequencing		1 mo
		1. DNA library preparation and indexing	\$3,000 (DNA extraction + library)	
Transcriptomics	A. Low-input RNA library sequencing	2. High-throughput sequencing	\$4,000 (sequencing 20 pooled samples)	2 mo
		1. RNA extraction and fragmentation	\$1,000 (RNA extraction + prelibrary processing)	
	B. Ultra-low-input RNA library sequencing	2. rRNA removal and cDNA conversion	\$2,000 (library)	
		3. Library preparation and indexing	\$4,000 (sequencing 20 pooled samples)	
	C. Targeted RNA sequencing	4. Sequencing, adapter removal		
		5. Alignment and mapping to genome		
Metabolomics and proteomics	LC-MS, LC-QTOF, NMR, GC-MS	1. Targeted amplicon or targeted enrichment		8-10 wk
		1. Liquid chromatography separation	\$12,000-\$150,000* (sample preparation, MS, analyses)	
		2. Electrospray and chemical ionization	\$10,000 (sample preparation, MS, analyses)	
Exposomics	MS measurements	3. Alignment of peaks based on m/z, analyses		8-10 wk
			\$150,000 (sample preparation, MS, analyses)	
Microbiomics	RNA-based 16S rRNA sequencing	1. Sequencing platform (see transcriptomics)	\$7,000 (sample and library, sequencing, analyses)	3-4 wk
	MS	1. Sequencing platform (see transcriptomics)	\$10,000 (sample and library, sequencing, analyses)	

ATAC-Seq, Assay for transposase accessible chromatin with high-throughput sequencing; *ChIP*, chromatin immunoprecipitation; *FFPE*, formalin-fixed, paraffin-embedded; *GC*, gas chromatography; *IP*, immunoprecipitation; *LC*, liquid chromatography; *NMR*, nuclear magnetic resonance; *QTOF*, Quadrupole Time-Of-Flight; *SNP*, single nucleotide polymorphism.

*Metabolomics pricing estimates are based on structured assay and service costs at the following institutes. Cost variations are dependent on assay type and number of metabolites needed for FA testing in the diverse clinical samples (as per September 20, 2017): <https://www.bcm.edu/research/advanced-technology-core-labs/lab-listing/metabolomics-core>; <https://genome.duke.edu/cores-and-services/proteomics-and-metabolomics/pricing>; http://www.bidmcmassspec.org/Service_Fees___Submission.html; <http://www.ucdenver.edu/academics/colleges/Graduate-School/academic-programs/biomol/program/strucbiocores/massspec/Pages/Metabolomics/Metabolomic-Services.aspx>; <https://www.bcm.edu/research/advanced-technology-core-labs/lab-listing/metabolomics-core>; <https://mass-spec.stanford.edu/rates>; and <http://depts.washington.edu/nwrmc/Assays%20and%20Fees>.

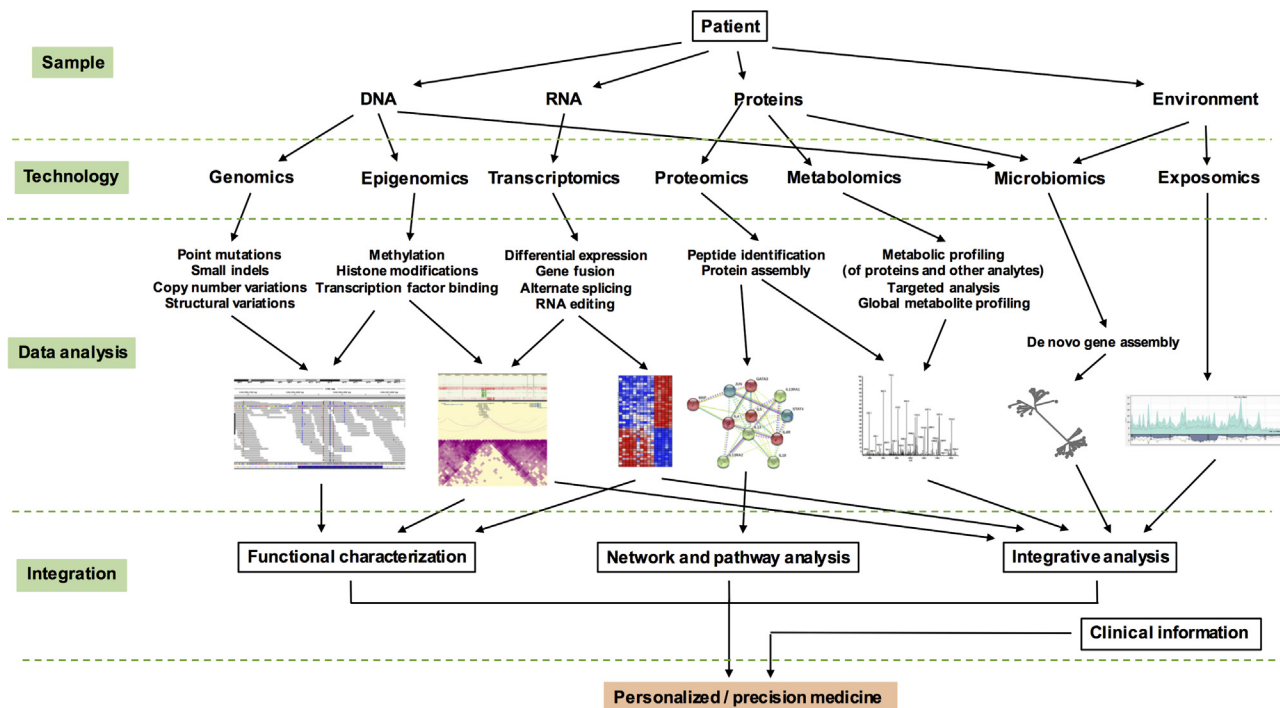


FIG 2. Schematic workflow for omics analyses: from samples to data generation by using omics technologies and to applications by using common data analyses approaches leading to personalized or precision medicine.

TABLE III. What do we know and what is still unknown?

What do we know?

- Common genetic variants are associated with FA and other atopic diseases.
- Methylation of *FOXP3* in regulatory T cells is directly linked to peanut sensitivity and tolerance.
- Bovine α -S1-casein peptide, which is found in maternal milk, causes allergic sensitization to cow's milk in breast-feeding infants.
- Metabolic fingerprints obtained from metabolomics studies might enable more specific and sensitive diagnostic biomarkers.
- Microbiome analysis of fecal samples indicates differences between the microbiota of IgE-mediated sensitized and nonallergic patients.
- Environmental exposures, including lifestyle, can alter the epigenome.
- Vast data from omics methods pose technical problems for downstream analysis and data interpretation.

What is still unknown?

- Desensitization resulting from OIT is often temporary, and the mechanism of recurrence of sensitization is unknown.
- Specific and sensitive biomarkers to distinguish severity of response on food challenge and diagnosis and prognosis (tolerance vs sustained unresponsiveness with OIT and natural history)
- Therapeutic strategies for enabling sustained unresponsiveness with immunotherapy
- The role and influence of allergens at a systems level
- Identification of gut microbes or their products that decrease or prevent allergic reactions
- The clinical value of detecting minute amounts of allergenic contaminants in food

relatively unbiased manner, the complement of proteins in serum. A more focused cell-based approach uses a panel of antibodies labeled with isotopically pure heavy metal isotope tags to probe

immune cells with a readout by using time-of-flight MS. This approach is known as mass cytometry or cytometry by time-of-flight.²⁵⁻²⁷ Conventional fluorescent dye-based flow cytometry can simultaneously analyze up to 18 proteins in a single cell, and further expansion in the number of parameters is limited by broad overlapping excitation and emission spectra. Cytometry by time-of-flight is a newer technique and allows for more than 40 targets (theoretically up to 100 isotopes) to be detected in a single sampling of cells with drastically less spillover between detector channels and lesser signal overlap.²⁸ In addition to global characterization of the immune system, this approach is being used currently to characterize allergen-specific T cells (by tracking their fate over the course of FA immunotherapy),^{28,29} the T-cell epitope repertoire,³⁰ and immunoglobulin neoantigens.³¹

Small-molecule *metabolomics* is the systematic identification, characterization, and quantification of all small metabolic products (low-molecular-weight compounds) created by using specific cellular processes in a biological system. There have been a number of studies demonstrating the usefulness of metabolomics (using breath condensate, urine, serum, and plasma) in patients with asthma and atopic dermatitis (AD).³² In children with peanut allergy, liquid chromatography MS analysis showed increased levels of uric acid in serum.³³ The authors hypothesize that uric acid likely activates dendritic cells by increasing CD86 and OX40 ligand expression. Comparison of plasma and saliva samples from the groups with peanut allergy and peanut tolerance with nuclear magnetic resonance spectroscopy found clear differences in metabolite levels between the 2 groups, and these differences might eventually enable the discovery of therapeutic targets or diagnostic biomarkers.^{34,35}

RNA level

Transcriptomics is the study of transcriptional profiles of a specific cell population or tissue by characterizing RNA molecules (generally mRNA but also transfer RNA and other noncoding RNA). Newer technologies, such as sequencing cDNA copies of RNA sequencing, are now used extensively to measure the transcriptome (transcriptionally active genes) of cells or organisms, although older technologies, such as microarrays (that can quantify a set of predetermined genes or sequences), are still in use. Although studies using RNA sequencing for FA are limited, transcriptional profiling with microarrays to analyze PBMCs of children with egg allergy (to unheated egg white protein) and children with egg tolerance (to baked egg) or nonallergic children showed increased expression of genes associated with allergic inflammation along with increased secretion of the cytokines IL-5, IL-9, and TNF- α .³⁶

Transcriptomic analysis has also been applied to understand allergens better and significant upregulation of genes associated with 3 pecan allergens (Car i 1, Car i 2, and Car i 4) was found in patients with pecan allergy.³⁷ In an effort to characterize wheat allergens, screening of a cDNA expression library of the wheat transcriptome with IgE antibodies from patients with wheat allergy revealed that the high-molecular-weight glutenin subunit Bx7 allergen contains repetitive IgE epitopes.³⁸

Recently, single-cell transcriptomic studies are beginning to explore the immune system in general³⁹ and the involvement of T_H1, T_H2,⁴⁰ and T_H17⁴¹ T cells in the pathogenesis of FA. RNA molecules with potential DNA regulatory functions, such as microRNAs, have been studied in patients with a variety of other disease states, but little is currently known about their potential roles in human FA.⁴²

DNA level

Study of the functions and interactions of all genes in the genome is termed *genomics*. Large-scale genome-wide association studies (GWASs) have helped to identify common genetic variants associated with diseases and have stimulated targeted drug therapy for diseases, such as rheumatoid arthritis and diabetes.⁴³ The potential utility of GWASs to advance our understanding of asthma and atopic diseases, such as allergic rhinitis, eosinophilic esophagitis, AD, and FA, is being recognized.⁴⁴

In patients with FA, a large GWAS of more than 2000 US subjects found 2 single nucleotide polymorphisms (rs7192 and rs9275596) associated with peanut allergy mapping to the HLA-DR and HLA-DQ regions at locus 6p21.32.⁴⁵ A smaller GWAS in an Australian population comparing 73 infants with challenge-proved peanut allergy with nonallergic control subjects similarly found an association between peanut allergy and HLA-DRB1, whereas HLA-DR and HLA-DQ alleles (HLA-DRB1*08 and DQB1*06:03P) were implicated with peanut allergy in another study.⁴⁷ Using a candidate gene approach, Hirota et al⁴⁸ found significant association between FA and 14 genetic loci from previous GWASs of AD and eosinophilic esophagitis, indicating the complex heterogeneous nature of these diseases and suggesting a common pathway among them.

Using high-resolution genome-wide single nucleotide polymorphism arrays, another study identified 2 potential gene candidates (*CTNNA3* and *RBFOX1*) harboring copy number variation loci that are enriched in pediatric patients with FA.⁴⁹ The authors confirmed the presence of these copy number

variations by means of quantitative PCR and found that knock-down of *CTNNA3* caused upregulation of CD63 and CD203c in mononuclear cells (that also included blood basophils) on stimulation with phorbol 12-myristate 13-acetate. Although the authors did not determine which cell types were most responsible for the upregulation of CD63 and CD203c in this setting, the findings suggest that *CTNNA3* can influence host responses to allergens.

Another important area in the study of FA is the somatic genomic rearrangements used in B and T cells to encode immunoglobulin and T-cell receptors (TCRs) for binding antigen because these are the molecules that determine the allergen specificity of the dysregulated adaptive immunity seen in patients with FA. Few allergen-specific antibodies or TCRs from allergic human patients have been fully defined at the level of their sequences, and the cellular pathways leading to isotype switching to IgE in human B cells are still under active study.⁵⁰⁻⁵⁴ Analysis of antibody and TCR sequences in FA research has benefitted greatly⁵⁵ from advances in DNA sequencing technologies, and several recent articles have identified and begun to characterize the features of these receptors in lymphocytes isolated from human blood, bone marrow, tonsils, and mucosal tissue sites in allergic patients.^{17,56-60} We believe that these studies represent the beginning of a full molecular characterization of the receptors responsible for specific FAs and might provide new biomarkers for the diagnosis and prognosis of allergic diseases, as well as for tracking responses to immunotherapies.

Epigenomics is the study of epigenetic changes that affect gene expression without modifying the DNA sequence. The most commonly studied epigenetic changes, which alter DNA accessibility and regulate gene expression, are DNA methylation (DNAm) and histone modifications. Hypermethylation of DNA is generally associated with gene silencing and genome stability, whereas hypomethylation typically results in the active transcriptional state of a gene.⁶¹ A phase I study by Syed et al⁶² found that DNAm of *FOXP3* in regulatory T cells was lower in participants with peanut allergy who underwent OIT compared with those who underwent standard care. Furthermore, they found that patients successfully treated with OIT but who lost their tolerance and regained sensitivity to peanuts after a period of peanut avoidance had increased methylation of *FOXP3* CpG sites.⁶² Similarly, in children with IgE-mediated cow's milk allergy, tolerance was associated with epigenetic regulation of the *FOXP3* gene.⁶³

A retrospective study that assessed approximately 450,000 CpGs in CD4⁺ T cells from a cohort of 12 children with IgE-mediated FA diagnosed at 12 months (and 12 nonallergic control subjects) identified 136 and 179 differentially methylated probes at birth and 12 months, respectively. Pathway analysis of these genes identified a number of mitogen-activated protein kinase signaling-associated molecules, which might contribute to suboptimal T-lymphocyte responses in early childhood and the development of FA.⁶⁴

Petrus et al⁶⁵ found a general increase in DNAm in children with challenge-proved cow's milk allergy compared with nonallergic children and varied methylation across regions of genes, such as *HTRA2*, which is known to be involved in immunologic pathways. Moreover, Martino et al²⁸ found a DNAm signature with 96 CpGs predictive for the reactivity of a child to multiple food allergens, which outperformed classical skin prick test and

allergen-specific IgE tests, suggesting epigenetic changes could potentially be used as a biomarker for predicting reactions on food challenges.

Environment level

In addition to genetic influences, nongenetic external exposures of a person, such as radiation, stress, lifestyle factors, microbial exposure, drugs, diet, pollution, socioeconomic and psychosocial influences, and other factors encountered during a patient's lifetime, also can influence health. The entire group of these nongenetic external exposures is termed the *exposome*, a word that was coined and defined by Dr Christopher Wild,^{66,67} and the analysis of the exposome is called "exposomics." The exposome (ie, the cumulative effects of chronic exposure from environmental factors) and genetic susceptibility together influence the development of atopic and other diseases.⁶⁸⁻⁷¹ Exposome study initiatives, such as HELIX (<http://www.projecthelix.eu>), EXPOsOMICS (<http://www.exposomicsproject.eu>), and HEALS (Health and Environment-Wide Association based on Large Scale population Surveys; <http://www.heals-eu.eu>), are in progress. A manually curated database of biomarkers of exposure to environmental risk factors (<http://exposome-explorer.iarc.fr>) and the Human Exposome Project databases (<http://humanexposomeproject.com>) are available as resources. Furthermore, the HERCULES Exposome Research Center provides valuable tools to assess the exposome (<http://emoryhercules.com>). Beyond the exposome, the importance of the metaexposome, which is the study of the bidirectional effect of the environment on human subjects and the human influence on all living systems and their genomes in human health, is being investigated.⁷²

Environmental allergen exposures in patients with AD with impaired skin barrier can result in allergic sensitization, including FA.⁷³ In the PEBBLES pilot study the prophylactic use of a ceramide-dominant emollient, "EpiCeram," from birth to 6 months of age has shown reduced incidence of AD and food sensitization at 12 months of age.⁷⁴ Other studies have linked FA with facial soap⁷⁵ and microbial exposure from dishwashing.⁷⁶ Early-life exposure to multiple allergenic foods in breastfeeding infants⁷⁷ and to peanuts⁷⁸⁻⁸¹ have shown reduced prevalence of FA.

Another major environmental exposure field is investigation of the microbiota, termed *microbiomics*. Our commensal microbiota outnumber the cells in our bodies by at least 10-fold,⁸² and it is now recognized that host-microbe interactions can have a profound effect on human health and disease.⁸³ Aberrations in the gut microbiome have been linked to diseases, such as diabetes and inflammatory bowel disease, as well as obesity.⁸⁴ Advances in our understanding of host-microbe interactions have been made possible by using new technologies, such as 16S rRNA sequencing, which permits precise identification of bacteria, including those that might not be grown in culture media. The American Gut Project initiative (www.Americangut.org)⁸⁵ aims to identify components contributing to the development and maintenance of an "ideal" gut flora and to determine how these are affected by diet and other lifestyle factors. Furthermore, a mechanistic framework has been proposed to characterize the ecology of the gut microbiome based on metabolic network modeling.⁸⁶

A number of studies have found differences in microbiota composition between patients with FA and nonallergic subjects,

with somewhat varying results. Analysis of fecal 16S rRNA sequence data from about 1800 participants with and without FAs found that adults with nut allergies have reduced Clostridiales, lower microbial diversity, and increased Bacteroidales in their gut microbiota.⁸² Another study found that resolution of milk allergy in children was associated with increases in Clostridia and Firmicutes.⁸³ Comparative analysis through parallel barcoded 454 pyrosequencing targeting the 16S rRNA gene hypervariable V3-V5 regions found increases in *Clostridium IV* and *Subdoligranulum* and decreases in *Bacteroides* and *Veillonella* species in children who were food sensitive compared with those without food sensitivities.⁸⁴ Microbial analysis of fecal samples of infants at the time of challenge-proved diagnosis of IgE-mediated allergy found that levels of *Clostridium sensu stricto* correlated with IgE levels.⁸⁵ In a longitudinal study of 12 food-sensitized and 154 nonallergic infants, a high ratio of Enterobacteriaceae/Bacteroidaceae in early infancy was associated with subsequent food sensitization.⁸⁷ Furthermore, Jones et al⁷³ have shown that in patients with AD dysfunction in the skin barrier associated with *Staphylococcus aureus* colonization can promote FA. These data suggest that microbial composition can be altered in patients with FA, but the nature of these changes and whether these changes precede allergy or follow onset of allergic disease awaits further study.

CHALLENGES IN DATA HANDLING AND PROCESSING

Researchers using multiple omics technologies face data-handling challenges because different technologies generate raw data in different formats. Even when using the same technology, there are differences in data content and formats because of variations in the age of the machine, instrumentation settings, or manufacturer's specifications. Therefore downstream analyses require a variety of computational and statistical methods; however, these pose challenges for standardized pipelines because changes in algorithm or assay characteristic can lead to different results.⁸⁸

Other challenges include the potentially extreme heterogeneity of biological data⁸⁹ and the sheer volume of data (terabytes or more) generated by using newer technologies. Large data sets pose storage, accessibility, and sharing problems and often require investigators to use high-performance computers or cloud-based analysis and data storage systems and maintain high security for protected health information.^{13,90} One should also be aware that the definition of what constitutes a "normal" finding is rapidly evolving as new data from larger and more diverse human populations are analyzed and curated.⁹¹

Dimensionality and missing values

One major consideration in "big data" research is the large number of features (dimensionality) across a relatively low number of biological samples. As data dimensionality increases, the noise in data also increases, creating unrelated or empty data points (sparsity). This form of data sparsity was originally conceptualized by Richard Bellman in 1957⁹² and expanded on in 2003.⁹³ Specialized dimensionality reduction methods are needed to handle statistical analysis of sparse data.⁹⁴

Another challenge in large data sets is the handling of missing data. Even if each measured variable has only a small number of

missing responses, in combination, the missing data can be numerous. In practice, a missing value is either removed from the analyses if the effect of the missing value is small or the missing values are filled in with statistically imputed values, but no single imputation approach is suitable for all circumstances.⁹⁵

Data integration (clinical and experimental)

Choosing an appropriate experimental or trial design is crucial for testing biological hypotheses and clearly determining clinical outcomes during an observational and/or after an interventional study. The involvement of statisticians or mathematicians is often critical for this step. Case-control studies are often used to address the heterogeneity associated with human diseases and are designed so that patients and control subjects are matched as closely as possible. However, because it is not possible to anticipate all potential confounding factors, case-control studies only address known features of the subjects and are therefore inherently less reliable than randomized double-blind studies. Integrative omics approaches attempt to use the interrogation of large data sets from sufficiently large numbers of participants and samples to tackle participant heterogeneity by incorporating the many sources of variability into statistical models.⁹⁶ However, even with increases in the features measured from each participant's samples and the large numbers of participants studied, the potential danger of unobserved confounding factors remains in any observational study design. In addition, the integration of clinical (questionnaires, symptoms, family history, and demographic) and experimental (laboratory investigations, biochemical assays, and tissue pathology) data sets poses a challenge for simultaneous consideration of effects from all factors, as well as appropriate quality assessment, verification, filtering, and scaling of data.⁹⁷

Accurately modeling the interactions between different layers of a biological system is a major concern for data integrative "systems biology" strategies. The corresponding relationships between different layers, such as DNAm and gene expression, microRNA expression and protein coding, and others, should be considered not only independently but also together during an integrative process.⁹⁸ In an attempt to model multiple features simultaneously and accurately, some new statistical models were developed. One among such promising models is the least absolute shrinkage and selection operator,⁹⁹ which has been used successfully for predicting asthma in preschool children¹⁰⁰ and lung function by using GWASs¹⁰¹ and that also can be applied in FA. In contrast to classical regression methods, the least absolute shrinkage and selection operator combines variable selection with regularization and derives easier interpretable statistical models. It uses continuous trajectories of regression coefficients as a penalty function for feature selection and is more stable and feasible for high dimensional data, such as those generated by application of omics approaches.⁹⁹ A recent review¹³ discusses several mathematic and statistical challenges in integrating omics data, and an article by Huang et al⁹⁸ summarizes both unsupervised and supervised computational methods for integration of multi-omics data.

FUTURE DIRECTIONS OF OMICS IN FA

FA is a complex and heterogeneous disease, and it is likely that multiple genetic variants play a role in the etiology of this

disorder. In addition to genetic and epigenetic factors,^{102,103} the microbiome¹⁰⁴ and the exposome¹⁰⁵ likely play critical roles in FA predisposition, disease progression, and the severity of reactions to offending foods. Omics technologies provide enormous quantities of data from different subjects and disease states, which enable the creation of detailed network maps representing physiologic pathways associated with health and disease.⁶⁹ To date in patients with FA, GWASs have uncovered some common genetic variants contributing to risk; however, additional sequencing studies targeting specific gene regions are needed to uncover other common genetic risk loci and rare genetic variants and to enable genotype-phenotype correlations.

A number of studies have indicated that *FOXP3* DNAm changes can influence FA or responses to treatment. Research into other epigenetic changes, including those mediated by histone modifications, might be useful in improving our ability to diagnose FA and to monitor its evolution or response to therapy. It is also likely that targeted alteration of the epigenome might soon be possible with advances in epigenome editing and DNazymes.¹⁰³ However, the risks and benefits of such approaches should be considered before they can be adopted for use in FA or other clinical settings.

Transcriptomics already has provided some information about differences in gene expression between allergic and nonallergic subjects and shows potential for identifying novel immune processes in FA. Proteomics could enable further progress in the identification and characterization of allergens and in the evaluation of food-processing techniques that might decrease the allergenic potential of allergic epitopes (likely by changing epitope confirmation), as well as identifying proteomic features associated with disease states or responses to therapy. The identification of metabolites synthesized by human or resident microbes could assist in molecular fingerprinting for improving FA prognosis, diagnosis, or targeted therapy. At present, omics technologies have generated a wealth of data and have assisted with the creation of molecular signatures, identification of disease susceptible genetic loci, and detection of potential biomarkers. Omics approaches have aided in further classification of heterogeneous diseases into endotypes and phenotypes. These cellular and molecular clues to disease need to be further tested in the laboratory and clinical trials before a precision medicine framework can be applied in FA. With these fast-growing efforts to generate and integrate data from different technologies, we think that the application of precision medicine and targeted therapy in patients with FA is likely within the next 10 years.

In summary, advances in both omics technologies and in statistical and computational techniques will permit the generation and analysis of rich sets of biological and environmental data, which can inform efforts to understand the cause and mechanisms of FA and assist in efforts to devise better ways to treat and perhaps even cure this disease.

REFERENCES

1. Panel NI-SE, Boyce JA, Assa'ad A, Burks AW, Jones SM, Sampson HA, et al. Guidelines for the diagnosis and management of food allergy in the United States: report of the NIAID-sponsored expert panel. *J Allergy Clin Immunol* 2010;126(suppl):S1-58.
2. Sicherer SH, Sampson HA. Food allergy: epidemiology, pathogenesis, diagnosis, and treatment. *J Allergy Clin Immunol* 2014;133:291-308.
3. Reisacher WR, Davison W. Immunotherapy for food allergy. *Curr Opin Otolaryngol Head Neck Surg* 2017;25:235-41.
4. Walkner M, Warren C, Gupta RS. Quality of life in food allergy patients and their families. *Pediatr Clin North Am* 2015;62:1453-61.

5. Turner PJ, Campbell DE. Epidemiology of severe anaphylaxis: can we use population-based data to understand anaphylaxis? *Curr Opin Allergy Clin Immunol* 2016;16:441-50.
6. Hill DA, Grundmeier RW, Ram G, Spergel JM. The epidemiologic characteristics of healthcare provider-diagnosed eczema, asthma, allergic rhinitis, and food allergy in children: a retrospective cohort study. *BMC Pediatr* 2016;16:133.
7. Andorf S, Borres MP, Block W, Tupa D, Bollyky JB, Sampath V, et al. Association of clinical reactivity with sensitization to allergen components in multifood-allergic children. *J Allergy Clin Immunol Pract* 2017;5:1325-34.e4.
8. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Food and Nutrition Board; Committee on Food Allergies: Global Burden, Causes, Treatment, Prevention, and Public Policy; Oria MP, Stallings VA, editors. Finding a Path to Safety in Food Allergy: Assessment of the Global Burden, Causes, Prevention, Management, and Public Policy. Washington (DC): National Academies Press (US); 2016. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK435937/>.
9. Patel CJ, Bhattacharya J, Butte AJ. An environment-wide association study (EWAS) on type 2 diabetes mellitus. *PLoS One* 2010;5:e10746.
10. Patel CJ, Chen R, Kodama K, Ioannidis JP, Butte AJ. Systematic identification of interaction effects between genome- and environment-wide associations in type 2 diabetes mellitus. *Hum Genet* 2015;132:495-508.
11. Kim DH, Kim YS, Son NI, Kang CK, Kim AR. Recent omics technologies and their emerging applications for personalised medicine. *IET Syst Biol* 2017;11: 87-98.
12. Tebani A, Afonso C, Marret S, Bekri S. Omics-based strategies in precision medicine: toward a paradigm shift in inborn errors of metabolism investigations. *Int J Mol Sci* 2016;17.
13. Alyass A, Turcotte M, Meyre D. From big data analysis to personalized medicine for all: challenges and opportunities. *BMC Med Genom* 2015;8:33.
14. Ebrahim A, Brunk E, Tan J, O'Brien EJ, Kim D, Szubin R, et al. Multi-omic data integration enables discovery of hidden biological regularities. *Nat Commun* 2016;7:13091.
15. Benson M. Clinical implications of omics and systems medicine: focus on predictive and individualized treatment. *J Intern Med* 2016;279:229-40.
16. Markopoulos C, van de Velde C, Zarca D, Ozmen V, Masetti R. Clinical evidence supporting genomic tests in early breast cancer: do all genomic tests provide the same information? *Eur J Surg Oncol* 2017;43:909-20.
17. Ryan JF, Hovde R, Glanville J, Lyu SC, Ji X, Gupta S, et al. Successful immunotherapy induces previously unidentified allergen-specific CD4+ T-cell subsets. *Proc Natl Acad Sci U S A* 2016;113:E1286-95.
18. Flocchi A, Wang J. -omic sciences: new horizons in food allergy. *Curr Opin Allergy Clin Immunol* 2015;15:234-6.
19. Yagami T, Haishima Y, Tsuchiya T, Tomitaka-Yagami A, Kano H, Matsunaga K. Proteomic analysis of putative latex allergens. *Int Arch Allergy Immunol* 2004; 135:3-11.
20. Nony E, Le Mignon M, Brier S, Martelet A, Moingeon P. Proteomics for allergy: from proteins to the patients. *Curr Allergy Asthma Rep* 2016;16:64.
21. Hirano K, Hino S, Oshima K, Nadano D, Urisu A, Takaiwa F, et al. Evaluation of allergenic potential for rice seed protein components utilizing a rice proteome database and an allergen database in combination with IgE-binding of recombinant proteins. *Biosci Biotechnol Biochem* 2016;80:564-73.
22. Bouakkadia H, Boutebba A, Haddad I, Vinh J, Guilloux L, Sutra JP, et al. [Immunoproteomics of non water-soluble allergens from 4 legumes flours: peanut, soybean, sesame and lentil]. *Ann Biol Clin (Paris)* 2015;73:690-704.
23. Candreva AM, Ferrer-Navarro M, Bronsoms S, Quiroga A, Curciarello R, Cauerhff A, et al. Identification of cross-reactive B-cell epitopes between Bos d 9.0101 (Bos Taurus) and Gly m 5.0101 (Glycine max) by epitope mapping MALDI-TOF MS. *Proteomics* 2017;17.
24. Orru S, Di Nicola P, Giuliani F, Fabris C, Conti A, Coscia A, et al. Detection of bovine alpha-S1-casein in term and preterm human colostrum with proteomic techniques. *Int J Immunopathol Pharmacol* 2013;26:435-44.
25. Bandyopadhyay S, Fisher DAC, Malkova O, Oh ST. Analysis of signaling networks at the single-cell level using mass cytometry. *Methods Mol Biol* 2017; 1636:371-92.
26. Bendall SC, Simonds EF, Qiu P, Amir el AD, Krutzik PO, Finck R, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* 2011;332:687-96.
27. Bandura DR, Baranov VI, Ornatsky OI, Antonov A, Kinach R, Lou X, et al. Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem* 2009; 81:6813-22.
28. Martino D, Allen K. Meeting the challenges of measuring human immune regulation. *J Immunol Methods* 2015;424:1-6.
29. Goswami R, Blazquez AB, Kosoy R, Rahman A, Nowak-Wegrzyn A, Berin MC. Systemic innate immune activation in food protein-induced enterocolitis syndrome. *J Allergy Clin Immunol* 2017;139:1885-96.e9.
30. Frazier A, Schulten V, Hinz D, Oseroff C, Sidney J, Peters B, et al. Allergy-associated T cell epitope repertoires are surprisingly diverse and include non-IgE reactive antigens. *World Allergy Org J* 2014;7:1-6.
31. Khodadoust MS, Olsson N, Wagar LE, Haabeth OA, Chen B, Swaminathan K, et al. Antigen presentation profiling reveals recognition of lymphoma immunoglobulin neoantigens. *Nature* 2017;543:723-7.
32. Scrivo R, Casadei L, Valerio M, Priori R, Valesini G, Manetti C. Metabolomics approach in allergic and rheumatic diseases. *Curr Allergy Asthma Rep* 2014; 14:445.
33. Kong J, Chalcraft K, Mandur TS, Jimenez-Saiz R, Walker TD, Goncharova S, et al. Comprehensive metabolomics identifies the alarmin uric acid as a critical signal for the induction of peanut allergy. *Allergy* 2015;70:495-505.
34. Peeters KA, Lamers RJ, Penninks AH, Knol EF, Bruijnzeel-Koomen CA, van Nesselrooij JH, et al. A search for biomarkers as diagnostic tools for food allergy: a pilot study in peanut-allergic patients. *Int Arch Allergy Immunol* 2011;155: 23-30.
35. Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Discov* 2016;15:473-84.
36. Kosoy R, Agashe C, Grishin A, Leung DY, Wood RA, Sicherer SH, et al. Transcriptional profiling of egg allergy and relationship to disease phenotype. *PLoS One* 2016;11:e0163831.
37. Mattison CP, Rai R, Settlege RE, Hinchliffe DJ, Madison C, Bland JM, et al. RNA-Seq analysis of developing pecan (*Carya illinoensis*) embryos reveals parallel expression patterns among allergen and lipid metabolism genes. *J Agric Food Chem* 2017;65:1443-55.
38. Baar A, Pahr S, Constantin C, Giavi S, Papadopoulos NG, Pelkonen AS, et al. The high molecular weight glutenin subunit Bx7 allergen from wheat contains repetitive IgE epitopes. *Allergy* 2014;69:1316-23.
39. Stubbington MJT, Rozenblatt-Rosen O, Regev A, Teichmann SA. Single-cell transcriptomics to explore the immune system in health and disease. *Science* 2017;358:58-63.
40. Stelmazczyk-Emmel A, Zawadzka-Krajewska A, Kopatys A, Demkow U. Th1, Th2, Th17, and regulatory cytokines in children with different clinical forms of allergy. *Adv Exp Med Biol* 2013;788:321-8.
41. Gaublomme JT, Yosef N, Lee Y, Gertner RS, Yang L, Wu C, et al. Single-cell genomics unveils critical regulators of Th17 cell pathogenicity. *Cell* 2015;163: 1400-12.
42. Bunning BJ, DeKruyff RH, Nadeau KC. Epigenetic changes during food-specific immunotherapy. *Curr Allergy Asthma Rep* 2016;16:87.
43. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 years of GWAS discovery: biology, function, and translation. *Am J Hum Genet* 2017;101:5-22.
44. Bunyavanich S, Schadt EE. Systems biology of asthma and allergic diseases: a multiscale approach. *J Allergy Clin Immunol* 2015;135:31-42.
45. Hong X, Hao K, Ladd-Acosta C, Hansen KD, Tsai H-J, Liu X, et al. Genome-wide association study identifies peanut allergy-specific loci and evidence of epigenetic mediation in US children. *Nat Commun* 2015;6:6304.
46. Martino DJ, Ashley S, Koplin J, Ellis J, Saffery R, Dharmage SC, et al. Genome-wide association study of peanut allergy reproduces association with amino acid polymorphisms in HLA-DRB1. *Clin Exp Allergy* 2017;47:217-23.
47. Madore AM, Vaillancourt VT, Asai Y, Alizadehfar R, Ben-Shoshan M, Michel DL, et al. HLA-DQB1*02 and DQB1*06:03P are associated with peanut allergy. *Eur J Hum Genet* 2013;21:1181-4.
48. Hirota T, Nakayama T, Sato S, Yanagida N, Matsui T, Sugiura S, et al. Association study of childhood food allergy with GWAS-discovered loci of atopic dermatitis and eosinophilic esophagitis. *J Allergy Clin Immunol* 2017;140:1713-6.
49. Li J, Fung I, Glessner JT, Pandey R, Wei Z, Bakay M, et al. Copy number variations in CTNNA3 and RFXO1 associate with pediatric food allergy. *J Immunology* 2015;195:1599-607.
50. Looney TJ, Lee JY, Roskin KM, Hoh RA, King J, Glanville J, et al. Human B-cell isotype switching origins of IgE. *J Allergy Clin Immunol* 2016;137:579-86.e7.
51. Davies JM, Platts-Mills TA, Aalberse RC. The enigma of IgE+ B-cell memory in human subjects. *J Allergy Clin Immunol* 2013;131:972-6.
52. Ramadani F, Upton N, Hobson P, Chan YC, Mzinza D, Bowen H, et al. Intrinsic properties of germinal center-derived B cells promote their enhanced class switching to IgE. *Allergy* 2015;70:1269-77.
53. Boyd SD, Hoh RA, Nadeau KC, Galli SJ. Immune monitoring for precision medicine in allergy and asthma. *Curr Opin Immunol* 2017;48:82-91.
54. Oettgen HC, Burton OT. IgE receptor signaling in food allergy pathogenesis. *Curr Opin Immunol* 2015;36:109-14.

55. Glanville J, Huang H, Nau A, Hatton O, Wagar LE, Rubelt F, et al. Identifying specificity groups in the T cell receptor repertoire. *Nature* 2017; 547:94-8.
56. Hoh RA, Joshi SA, Liu Y, Wang C, Roskin KM, Lee J-Y, et al. Single B-cell deconvolution of peanut-specific antibody responses in allergic patients. *J Allergy Clin Immunol* 2016;137:157-67.
57. Patil SU, Ogunniyi AO, Calatroni A, Tadigotla VR, Ruiter B, Ma A, et al. Peanut oral immunotherapy transiently expands circulating Ara h 2-specific B cells with a homologous repertoire in unrelated subjects. *J Allergy Clin Immunol* 2015;136: 125-34.e12.
58. Levin M, Levander F, Palmason R, Greiff L, Ohlin M. Antibody-encoding repertoires of bone marrow and peripheral blood—a focus on IgE. *J Allergy Clin Immunol* 2017;139:1026-30.
59. Levin M, King JJ, Glanville J, Jackson KJ, Looney TJ, Hoh RA, et al. Persistence and evolution of allergen-specific IgE repertoires during subcutaneous specific immunotherapy. *J Allergy Clin Immunol* 2016;137:1535-44.
60. Ramadani F, Bowen H, Upton N, Hobson PS, Chan YC, Chen JB, et al. Ontogeny of human IgE-expressing B cells and plasma cells. *Allergy* 2017;72:66-76.
61. Deaton AM, Bird A. CpG islands and the regulation of transcription. *Genes Dev* 2011;25:1010-22.
62. Syed A, Garcia MA, Lyu SC, Bucayu R, Kohli A, Ishida S, et al. Peanut oral immunotherapy results in increased antigen-induced regulatory T-cell function and hypomethylation of forkhead box protein 3 (FOXP3). *J Allergy Clin Immunol* 2014;133:500-10.
63. Paparo L, Nocerino R, Cosenza L, Aitoro R, D'Argenio V, Del Monaco V, et al. Epigenetic features of FoxP3 in children with cow's milk allergy. *Clin Epigenet* 2016;8:86.
64. Martino D, Joo JE, Sexton-Oates A, Dang T, Allen K, Saffery R, et al. Epigenome-wide association study reveals longitudinally stable DNA methylation differences in CD4⁺ T cells from children with IgE-mediated food allergy. *Epigenetics* 2014;9:998-1006.
65. Petrus NC, Henneman P, Venema A, Mul A, van Sinderen F, Haagsmans M, et al. Cow's milk allergy in Dutch children: an epigenetic pilot survey. *Clin Transl Allergy* 2016;6:16.
66. Wild CP. Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 2005;14:1847-50.
67. Wild CP. The exposome: from concept to utility. *Int J Epidemiol* 2012;41:24-32.
68. Hua X, Goedert JJ, Pu A, Yu G, Shi J. Allergy associations with the adult fecal microbiota: analysis of the American Gut Project. *EBioMedicine* 2016; 3:172-9.
69. Bunyavanich S, Shen N, Grishin A, Wood R, Burks W, Dawson P, et al. Early-life gut microbiome composition and milk allergy resolution. *J Allergy Clin Immunol* 2016;138:1122-30.
70. Chen CC, Chen KJ, Kong MS, Chang HJ, Huang JL. Alterations in the gut microbiotas of children with food sensitization in early life. *Pediatr Allergy Immunol* 2016;27:254-62.
71. Ling Z, Li Z, Liu X, Cheng Y, Luo Y, Tong X, et al. Altered fecal microbiota composition associated with food allergy in infants. *Appl Environ Microbiol* 2014;80:2546-54.
72. Renz H, Holt PG, Inouye M, Logan AC, Prescott SL, Sly PD. An exposome perspective: early-life events and immune development in a changing world. *J Allergy Clin Immunol* 2017;140:24-40.
73. Jones AL, Curran-Everett D, Leung DYM. Food allergy is associated with *Staphylococcus aureus* colonization in children with atopic dermatitis. *J Allergy Clin Immunol* 2016;137:1247-8.
74. Lowe A. A randomized trial of a barrier lipid replacement strategy for the prevention of atopic dermatitis and allergic sensitisation: the PEBBLES pilot study. *J Invest Dermatol* 2017;137:S72.
75. Hiragun M, Ishii K, Yanase Y, Hiragun T, Hide M. Remission rate of patients with wheat allergy sensitized to hydrolyzed wheat protein in facial soap. *Allergol Int* 2016;65:109-11.
76. Hesselmar B, Hicke-Roberts A, Wennergren G. Allergy in children in hand versus machine dishwashing. *Pediatrics* 2015;135:E590-7.
77. Perkin MR, Logan K, Tseng A, Raji B, Ayis S, Peacock J, et al. Randomized trial of introduction of allergenic foods in breast-fed infants. *N Engl J Med* 2016;374: 1733-43.
78. Turcanu V, Brough HA, Du Toit G, Foong RX, Marrs T, Santos AF, et al. Immune mechanisms of food allergy and its prevention by early intervention. *Curr Opin Immunol* 2017;48:92-8.
79. Greenhawt M. The Learning Early About Peanut Allergy Study: the benefits of early peanut introduction, and a new horizon in fighting the food allergy epidemic. *Pediatr Clin North Am* 2015;62:1509.
80. Greenhawt M, Fleischer DM, Chan ES, Venter C, Stukus D, Gupta R, et al. LEAPing through the looking glass: secondary analysis of the effect of skin test size and age of introduction on peanut tolerance after early peanut introduction. *Allergy* 2017;72:1254-60.
81. Allen KJ, Koplin JJ. Does LEAP change the screening paradigm for food allergy in infants with eczema? *Clin Exp Allergy* 2016;46:42-7.
82. Spasova DS, Surh CD. Blowing on embers: commensal microbiota and our immune system. *Front Immunol* 2014;5:318.
83. Rachid R, Chatila TA. The role of the gut microbiota in food allergy. *Curr Opin Pediatr* 2016;28:748-53.
84. Blum HE. The human microbiome. *Adv Med Sci* 2017;62:414-20.
85. McDonald D, Birmingham A, Knight R. Context and the human microbiome. *Microbiome* 2015;3:52.
86. Sung J, Hale V, Merkel AC, Kim P-J, Chia N. Metabolic modeling with big data and the gut microbiome. *Appl Transl Genom* 2016;10:10-5.
87. Azad MB, Konya T, Guttman DS, Field CJ, Sears MR, HayGlass KT, et al. Infant gut microbiota and food sensitization: associations in the first year of life. *Clin Exp Allergy* 2015;45:632-43.
88. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 2014;15:819-912.
89. Marx V. Biology: the big challenges of big data. *Nature* 2013;498:255-60.
90. Schadt EE, Linderman MD, Sorenson J, Lee L, Nolan GP. Cloud and heterogeneous computing solutions exist today for the emerging big data problems in biology. *Nat Rev Genet* 2011;12:224.
91. Olson MV. Genome-sequencing anniversary. What does a “normal” human genome look like? *Science* 2011;331:872.
92. Bellman R. Dynamic programming. Princeton (NJ): Princeton University Press; 1957.
93. Bellman R. Dynamic Programming. Mineola (NY): Dover Publications; 2003.
94. Binder H, Benner A, Bullinger L, Schumacher M. Tailoring sparse multivariable regression techniques for prognostic single-nucleotide polymorphism signatures. *Stat Med* 2013;32:1778-91.
95. Webb-Robertson B-JM, Wiberg HK, Matzke MM, Brown JN, Wang J, McDermott JE, et al. Review, evaluation, and discussion of the challenges of missing value imputation for mass spectrometry-based label-free global proteomics. *J Proteome Res* 2015;14:1993-2001.
96. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol* 2017;18:83.
97. Xu C, Zhang J-G, Lin D, Zhang L, Shen H, Deng H-W. A systemic analysis of transcriptomic and epigenomic data to reveal regulation patterns for complex disease. *G3 (Bethesda)* 2017;7:2271-9.
98. Huang S, Chaudhary K, Garmire LX. More is better: recent progress in multi-omics data integration methods. *Front Genet* 2017;8:1005-12.
99. Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc Series B Stat Methodol* 1996;58:267-88.
100. Pescatore AM, Spycher BD, Jurca M, Gaillard EA, Kuehni CE. Environmental and socioeconomic data do not improve the Predicting Asthma Risk in Children (PARC) tool. *J Allergy Clin Immunol* 2015;136:1395-7.
101. Yao TC, Du G, Han L, Sun Y, Hu D, Yang JJ, et al. Genome-wide association study of lung function phenotypes in a founder population. *J Allergy Clin Immunol* 2014;133:248-55, e1-10.
102. Li J, Maggadottir SM, Hakonarson H. Are genetic tests informative in predicting food allergy? *Curr Opin Allergy Clin Immunol* 2016;16:257-64.
103. Potaczek DP, Harb H, Michel S, Alhamwe BA, Renz H, Tost J. Epigenetics and allergy: from basic mechanisms to clinical applications. *Epigenomics* 2017;9: 539-71.
104. Hong SW, Kim KS, Surh CD. Beyond hygiene: commensal microbiota and allergic diseases. *Immune Netw* 2017;17:48-59.
105. Feldweg AM. Food-dependent, exercise-induced anaphylaxis: diagnosis and management in the outpatient setting. *J Allergy Clin Immunol Pract* 2017;5:283-8.