

Gene-environment interactions in asthma and allergic diseases: Challenges and perspectives

Francine Kauffmann, MD,^{a,b} and Florence Demenais, MD^{c,d,e} *Villejuif and Paris, France*

INFORMATION FOR CATEGORY 1 CME CREDIT

Credit can now be obtained, free for a limited time, by reading the review articles in this issue. Please note the following instructions.

Method of Physician Participation in Learning Process: The core material for these activities can be read in this issue of the Journal or online at the JACI Web site: www.jacionline.org. The accompanying tests may only be submitted online at www.jacionline.org. Fax or other copies will not be accepted.

Date of Original Release: December 2012. Credit may be obtained for these courses until November 30, 2014.

Copyright Statement: Copyright © 2012-2014. All rights reserved.

Overall Purpose/Goal: To provide excellent reviews on key aspects of allergic disease to those who research, treat, or manage allergic disease.

Target Audience: Physicians and researchers within the field of allergic disease.

Accreditation/Provider Statements and Credit Designation: The American Academy of Allergy, Asthma & Immunology (AAAAI) is accredited by the Accreditation Council for Continuing Medical Education (ACCME) to provide continuing medical education for physicians. The AAAAI designates this journal-based CME activity for a maximum of

1 AMA PRA Category 1 Credit™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

List of Design Committee Members: Francine Kauffmann, MD, and Florence Demenais, MD

Activity Objectives

1. To learn how most of the validated gene-environment interactions in asthma were identified.
2. To understand what a gene-environment-wide interaction study is and why these types of studies are emerging.
3. To learn about the challenges in detecting gene-environment interactions and the possible solutions to overcome challenges and improve gene-environment interaction discovery.
4. To understand why the study of gene-environment interactions is useful.

Recognition of Commercial Support: This CME activity has not received external commercial support.

Disclosure of Significant Relationships with Relevant Commercial

Companies/Organizations: F. Kauffmann and F. Demenais have received research support from the ANR French National Agency of Research.

The concept of gene-environment (GxE) interactions has dramatically evolved in the last century and has now become a central theme in studies that assess the causes of human disease. Despite the numerous efforts to discover genes associated in asthma and allergy through various approaches, including the recent genome-wide association studies, investigation of GxE interactions has been mainly limited to candidate genes, candidate environmental exposures, or both. This review discusses the various strategies from hypothesis-driven strategies to the full agnostic search of GxE interactions with an illustration from recently published articles. Challenges raised

by each piece of the puzzle (ie, phenotype, environment, gene, and analysis of GxE interaction) are put forward, and tentative solutions are proposed. New perspectives to integrate various types of data generated by new sequencing technologies and to progress toward a systems biology approach of disease are outlined. The future of a molecular network-based approach of disease to which GxE interactions are related requires space for innovative and multidisciplinary research. Assembling the various parts of a puzzle in a complex system could well occur in a way that might not necessarily follow the rules of logic. (J Allergy Clin Immunol 2012;130:1229-40.)

Key words: Gene, environment, gene-environment interaction, asthma, allergy

From ^aINSERM, CESP Centre for research in Epidemiology and Population Health, U1018, Respiratory and Environmental Epidemiology Team, Villejuif; ^bUniversité Paris Sud, UMRS 1018, Villejuif; ^cINSERM, U946, Genetic Variation and Human Diseases Unit, Paris; ^dUniversité Paris Diderot, Sorbonne Paris Cité, Institut Universitaire d'Hématologie, Paris; and ^eFondation Jean Dausset-Centre d'Etude du Polymorphisme Humain (CEPH), Paris.

Supported by grants from the French National Agency for Research (ANR-2010-PRSP-003 IAGO and ANR-11-BSV1-027-01-GWIS-AM).

Received for publication October 2, 2012; revised October 27, 2012; accepted for publication October 30, 2012.

Corresponding author: Florence Demenais, MD, INSERM UMR-946, Fondation Jean Dausset-CEPH, 27 rue Juliette Dodu, F-75010 Paris, France. E-mail: florence.demenais@inserm.fr.

0091-6749/\$36.00

© 2012 American Academy of Allergy, Asthma & Immunology

<http://dx.doi.org/10.1016/j.jaci.2012.10.038>

Asthma and allergy, as most common diseases, are believed to result from the combined effect of genes and environmental factors and their interactions. The nature-nurture debate finds its foundations in the 19th century with the work of 2 cousins: Charles Darwin, the father of the theory of evolution, and Francis Galton, the father of human behavioral genetics. The idea of a potential role of gene-environment (GxE) interaction in respiratory disease can be traced back to as early as 1938, when Haldane¹ suggested that genetic differences might explain the variability in the responses of potters to industrial exposures with respect to the occurrence of bronchitis and death. The concept of GxE

Abbreviations used

CNV:	Copy number variant
ETS:	Environmental tobacco smoke
GEWIS:	Gene-environment-wide interaction study
GWAS:	Genome-wide association study
GxE:	Gene-environment
SNP:	Single nucleotide polymorphism

interactions has dramatically evolved in the last century and has now become a central theme in studies that assess the causes of human disease.^{2,3} The way of addressing the issue of GxE interaction has been markedly driven by the specialty of researchers interested in the topic (Fig 1). Disentangling complex issues by defining successive steps focused on simpler aspects has long been the basis of science since Descartes in the 17th century.⁴ However, disciplines vary according to their conceptualization of causation of disease.

For genetics-oriented scientists, the basis of the determinants of disease was seen as starting with genes, followed by gene products influencing biological and physiologic phenotypes up to the occurrence of clinical symptoms, as presented by Schork⁵ in the Transatlantic Conference on asthma genetics held in 1997, the beginning of the genome-wide linkage screens era. The environment was acknowledged as a question to be addressed after genetic aspects had been addressed, with a marginal role for the most “pure” biological steps. That frame might correspond to the succession of events going from genetic polymorphisms within the *IL4* or *IL13* genes to the T_H2-oriented immune response to IgE levels to airway hyperresponsiveness and finally to asthma symptoms.⁶ By contrast, the view of environmentalists,⁷ also processing by successive steps, was merely to consider genetic susceptibility. An example is occupational asthma, for which the environment is well defined; the internal dose will depend on the use of a mask at work for occupational exposure to dust, gases, and fumes or could be assessed based on cotinine levels for environmental tobacco smoke (ETS) exposure; and altered structure function might result from bronchial hyperresponsiveness. “Susceptibility” is mentioned in the classical framework regarding the effect of exposure up to manifestation of disease (Fig 1), which is still in use. The consideration of genetic modifiers in the analysis of environmental determinants of disease has been viewed sometimes as potentially decreasing the role of social control on some factors and having the potential of discrimination, especially in the field of occupational exposures.⁸ Such ethical concerns, often not explicit, have led to some reluctance to study the genetic components of disease. More research is needed on the ethical aspects of GxE interactions, especially regarding the formulation of research questions and the interpretation of findings.⁹

Technological advances in large-scale biology have greatly influenced the recent evolution of the field.^{10,11} There have been a number of initiatives worldwide to accelerate investigation of GxE interplay in health and disease, among which was the Genes, Environment and Health Initiative launched by the National Institutes of Health in 2006,¹² which had the primary goals of identifying novel associations of genetic variants with disease, improving measures of environmental factors through the development of exposure biology tools, and developing new methods to identify interactions between genes and environmental

exposures. Although the first 2 goals have been addressed separately, a comprehensive approach to understanding the GxE interactions underlying complex diseases still faces numerous challenges. Workshops gathering experts from various disciplines have been recently organized to propose recommendations about the future direction of GxE interaction studies.^{13,14}

The term “gene-environment interaction” has different meanings, depending on the context. A statistical interaction represents any departure from a pure main effects model. Although qualitative interactions require interaction terms on any scale of measurement, quantitative interactions are scale dependent.¹⁵ Statistical interaction differs from biological interaction. Biological interaction refers to the dependence of the effect of one factor at the cellular or molecular level on the presence or absence of the other, whereas the public health concept of synergy means that the joint effect of 2 factors is greater than the sum of the excess risks from each factor.¹⁵ Fig 2 illustrates theoretical models of relationships between genotype and environment,¹⁶ with examples in the field of asthma.¹⁷

The study of GxE interaction is useful for several reasons. Among the many advantages of studying gene interactions in asthma and allergy is the opportunity to identify novel genes and novel environmental risk factors that have an effect in subgroups of subjects (environmentally exposed for the former or genetically susceptible for the latter) and to provide better insight into the complex biological mechanisms and pathways underlying disease. Also, characterization of the increased risk of asthma associated with the joint effects of genes and the environment is of public health relevance given the dramatic increase in the prevalence of asthma and allergic diseases over the last decades in developed countries and more recently in developing countries.¹⁸ Studies of the effect of genes in populations sharing the same genetic background but living in different environments and lifestyles can be particularly instructive.^{19,20} Finally, investigation of GxE interactions might have a translational value in the field of pharmacogenetics, in which the genetic control of response to drugs can be viewed as a type of GxE interaction. However, to date, small numbers of genes involved in asthma-related drug responses have been identified.²¹

WHAT HAS BEEN DONE AND WHERE ARE WE NOW IN GxE INTERACTIONS FOR ASTHMA?

A number of strategies can be used to identify GxE interactions that depend on the level of knowledge²² one can have on the interaction itself and on each of its components: genes, the environment, or both (Table I).^{17,21,23-45} The first and simplest situation is that of a candidate interaction, when both the gene and the environment are known and assumed to be involved in a direct physical or chemical reaction. The most well-known and extensively studied candidate interaction in the asthma literature is the interaction between the *CD14* gene, which encodes the endotoxin receptor, an essential component of the gram-negative bacteria wall, and environmental exposure to endotoxin.^{23,24} Other examples of candidate interactions are those of *HLA* genes with allergens²⁵ that are presented to T lymphocytes by HLA molecules and that of Toll-like receptor genes with infectious agents²⁶ that are first recognized by Toll-like receptors to activate the innate immune response. Other candidate interactions correspond to lower levels of knowledge than mechanistic interactions, including knowledge on immunologic or functional mechanisms or

The third (principle), to conduct my thoughts in such order that, by commencing with objects the simplest and easiest to know, I might ascend by little and little, and, as it were, step by step, to the knowledge of the more complex; assigning in thought a certain order even to those objects which in their own nature do not stand in a relation of antecedence and sequence.

The diagram illustrates the progression from a gene to a disease, showing the relative contributions of genetic and environmental factors at each stage. A vertical DNA sequence (A, C, G, A, G, T, A, C, C, G, A, A, T, I, G, G, T, C, T, G, T, A, A) is shown on the left. The main flow is represented by a series of arrows and boxes:

- Gene** (blue box) → **90%** → **Gene product** (black box)
- Gene product** → **75%** → **Low level physiological phenotype** (black box)
- Low level physiological phenotype** → **50%** → **Intermediate phenotype** (black box)
- Intermediate phenotype** → **25%** → **High level clinically relevant phenotype** (black box)
- High level clinically relevant phenotype** → **5%** → **Disease (Clinical signs)** (black box)

A green bar at the bottom, labeled **Environmental contribution**, indicates the increasing influence of environmental factors as the phenotype progresses from low level to high level.

```

graph LR
    E[Exposure (E)] --> ID[Internal dose (ID)]
    ID --> BED[Biologically effective dose (BED)]
    BED --> EBE[Early biological effect (EBE)]
    EBE --> ASF[Altered structure/function (ASF)]
    ASF --> CD[Clinical disease (CD)]
    CD --> PS[Prognostic significance (PS)]
    S[Susceptibility] -.-> ID
    S -.-> BED
    S -.-> EBE
    S -.-> ASF
    S -.-> CD
    S -.-> PS
  
```

With the recent development of high-throughput genotyping technologies (based on Illumina [San Diego, Calif] or Affymetrix [Santa Clara, Calif] single nucleotide polymorphism [SNP] arrays), investigation of GxE interactions at the genome-wide level (gene-environment-wide interaction studies [GEWISs]) are emerging and receiving growing interest.^{15,52,53} One way to simplify the investigation of GxE interactions at the genome-wide level when the environment is known is to conduct a GWAS of environment-related subphenotypes. This was illustrated by the first GWAS of toluene diisocyanate-induced asthma, which, although underpowered, suggested a potential association with the catenin (cadherin-associated protein) $\alpha 3$ (*CTNNA3*) gene.³⁸ This observation, as with many others, has not been yet replicated. Other examples in which such approaches have been attempted include GWASs of aspirin-induced asthma³⁹ and of asthma phenotypes related to drug response.²¹ GWASs of various types of occupational

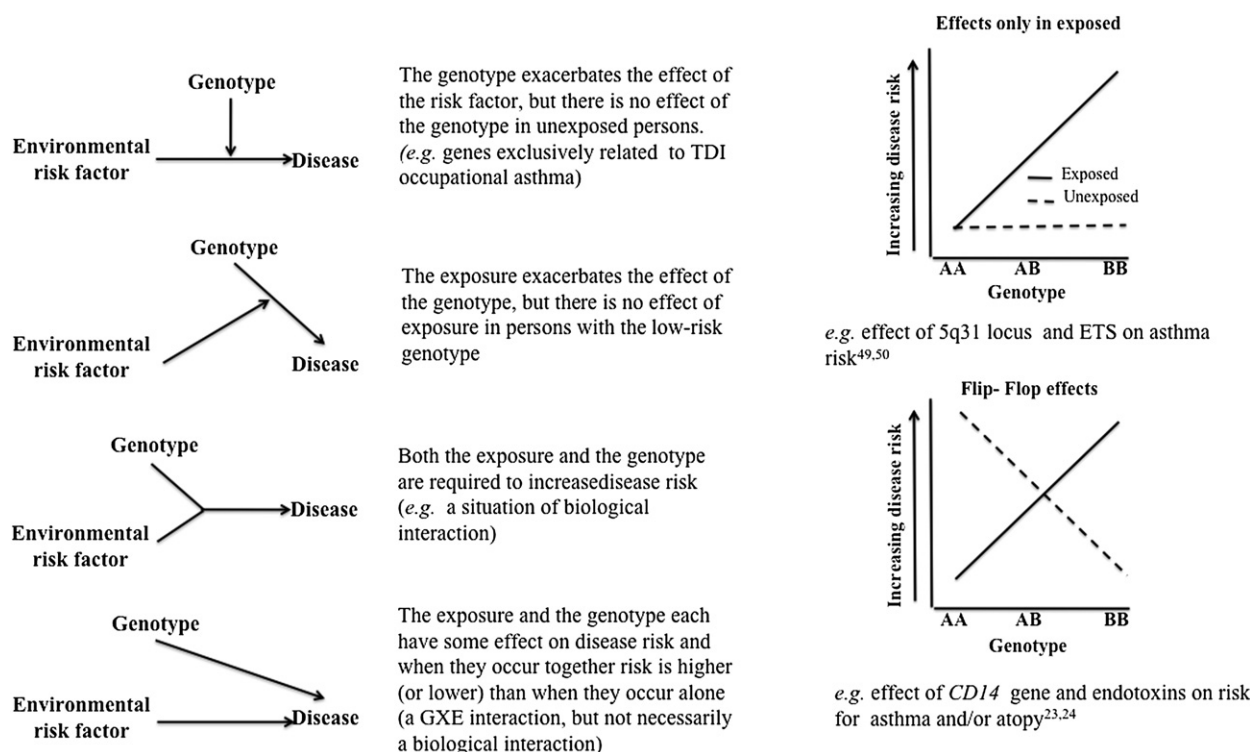


FIG 2. Models of GxE interaction and examples in the area of asthma. The *left panel* presents theoretical models of the effects of high-risk genotype and environmental exposure on disease (modified from Ottman¹⁶). The *right panel* presents 2 situations taken from the asthma literature in which only the second (flip-flop) is independent of the scale of measurement (modified from Ober and Vercelli¹⁷).

asthma can be similarly conducted but face the difficulty of replication if the discovered genes are specific to a given exposure.

The first published genome-wide environment interaction study for asthma explored interactions between farm-related exposure and genome-wide SNPs in 1700 children from 4 rural regions of Central Europe.⁴⁰ This well-designed study did not reveal any significant interactions with common SNPs (interacting allele frequencies from 30% to 70%) for which this study was well powered or with SNPs previously published for interaction with asthma. Nevertheless, interactions were detected with rarer SNPs in genes that merit attention, such as the glutamate receptor pathway, which is shown to interact with coffee (a well-known bronchodilator) through a GEWIS of Parkinson disease.⁵⁴ The authors were cautious regarding these new observations and acknowledge the need of replication in independent farm populations.⁴⁰ Other GEWISs include the investigation of genome-wide interactions of SNPs with smoking exposure in asthma.⁴¹

Finally, the most complex situation is when both the gene and the environment are unknown. Although the search for GxE interactions in that context has not yet been undertaken, this raises the issue of defining the exposome (large sets of environmental exposures or all putative environmental determinants of asthma)⁴² and of reducing the problem by restricting the study to specific asthma phenotypes.⁵⁵

Therefore all these previous studies showed that most of the well-validated GxE interactions were found by using knowledge of biological processes, the effect of environmental exposures, and/or pathophysiologic mechanisms (subtypes of disease, sometimes called endotypes).⁶ Agnostic approaches of GxE interactions are still in their infancy, but methodological research for proposing both optimal study design and analytic methods is being actively pursued.^{13-15,56}

WHAT ARE THE CHALLENGES TO DETECT GxE INTERACTIONS, AND WHICH TENTATIVE SOLUTIONS CAN BE PROPOSED?

In an attempt to illustrate the current issues in an unbiased way, we performed a literature search for articles published in 2011 or 2012, combining 2 bibliographic sources: (1) Hugenet⁵⁷ (with “asthma” as the key word in the curated category “gene-environment interaction”), which retrieved 277 articles published since 2001, 80 of which appeared in 2011-2012, and (2) PubMed (search with “asthma AND gene AND environment AND interaction”), which retrieved 150 articles, 49 of which were published in 2011-2012. Table II^{20,40,43,44,58-64} summarizes key aspects of the 11 articles present in both sources. These 11 articles cover all continents and were published in 10 different journals. Except for 1 article, the number of subjects was usually at least several hundred, with a few studies including more than 1000 subjects with asthma. Environmental determinants studied were diverse and mostly concerned smoking, farming, indoor environment, occupational environment, air pollution, and the macroenvironment. A variety of rationales were given to undertake interaction analysis from candidate interactions to relevance of marginal effects and finally to agnostic search for a specific environment. None of the 11 articles showed a very clear interaction and had varying interpretations from the authors from classical enthusiasm that can sometimes be viewed as overinterpretation to extreme caution that can be viewed as underinterpretation.

The challenges are well illustrated by this sample of 11 articles, but identifying the issues raised by GxE interactions in a more systematic manner and proposing tentative solutions can provide new opportunities to improve GxE discovery (Table III). We

TABLE I. Gradient of *a priori* hypotheses for testing GxE interactions: from candidate interaction to full agnostic search

Strategy according to hypotheses on G, E, and/or GxE interaction	Examples
A: Hypothesis driven, using previous knowledge on both G and E	
Candidate interaction based on a specific hypothesis regarding GxE interactions (both known): mechanical interaction or based on immunologic or functional knowledge or involving a specific environment-associated phenotype	<i>CD14</i> (endotoxin receptor) with exposure to endotoxin ^{23,24} <i>HLA</i> with any allergen ²⁵ Innate immunity genes with infectious agents ²⁶ Candidate genes for aspirin or TDI-induced asthma <i>ADRB2</i> with β_2 -agonist response
Known marginal effects of both G and E on disease without a specific hypothesis regarding GxE interaction	ETS in childhood and 17q21 variants; candidate interaction with respect to the same subphenotype (early-onset asthma) ²⁷ Genes of interest for asthma, such as <i>IL4</i> , with molds ⁴³
Pathway approach: testing interaction between a set of genetic variants and a given environment	ETS, air pollution, farming, and occupational exposure with sets of potential candidate genes (a few genes [<10] or many genes [hundreds of genes]) belonging to pathways defined by a systematic search using bioinformatic tools ^{17,28-37}
B: Agnostic for G	
Genome-wide interaction investigation of GxE interactions for a given environment	GWAS of environment-related phenotype (TDI asthma, ³⁸ aspirin-related asthma, ³⁹ specific sensitization, GWAS of drug response-related phenotype to a given treatment [pharmacogenomics]) ²¹ Farming exposure with genome-wide SNPs (GEWIS) ⁴⁰ Smoking exposure with genome-wide SNPs (GEWIS) ⁴¹
C: Agnostic for G and E	
Full agnostic search: testing all genes with all environments (hundreds of environmental determinants or all putative environmental determinants of asthma with all genes)	Not yet done (issue of defining exposome); data-reduction approach ^{42,44,45}

ADRB2, Adrenoceptor β_2 gene; *E*, environment; *G*, gene; *TDI*, toluene diisocyanate.

present these challenges and possible solutions successively for the phenotype under study, environmental exposures, genes, and finally GxE interactions.

The definition of the outcome (phenotype) under study is of major importance. It is well known that asthma is a heterogeneous disease and that there is not one asthma but several asthmas.⁵⁵ It has been shown that the number of SNPs found to be associated robustly with a complex disease by using GWASs was higher for well-defined diseases (eg, Crohn disease) than for diseases with less precise definitions and affecting multiple organ systems (eg, cardiovascular diseases).¹⁴ There are various ways of improving the phenotype definition. One way is to consider physiologic or biological phenotypes (endophenotypes) related to the disease process. Many phenotypes involved in allergy, inflammation, and airway remodeling are associated with asthma. Although genetic and environmental determinants of these phenotypes have been explored separately, studies of GxE interactions have been limited up to now and need to be expanded. Attempts have been made to define subphenotypes of asthma that were initially based on combinations of clinical characteristics but now tend to include biology.⁵⁵ Examples of these subphenotypes are early-onset allergic asthma related to T_H2 cytokines, T_H2 -related genes and the 17q12-21 locus, obesity-related asthma, and involvement of the oxidative stress pathway.⁵⁵ Use of unbiased and statistically based approaches, such as cluster analysis or latent-class analysis, are appealing because they allow one to define new phenotypes from a number of features related to disease occurrence, evolution, therapy, and biological and physiologic disease-related phenotypes.⁶⁵⁻⁶⁷ One component that is known to play a major role in the pathophysiology of asthma and has been highlighted by clustering-based approaches is time. Therefore the 2-dimensional problem of GxE interaction can be extended to a 3-dimensional problem (gene-environment-time interaction, Fig 3). Taking time of asthma onset into account can be instrumental to characterize novel GxE interactions.²⁷ As an alternative to stratify

on age of disease onset, time to disease onset can be the outcome under study, as in a recent GWAS of age of onset of asthma.⁶⁸ More precise phenotyping can also be obtained from longitudinal studies that provide not only information on the evolution of disease but also on repeated measures of disease-related phenotypes. Although most genetic studies have examined single phenotypes (univariate analysis), joint analysis of multiple phenotypes through multidimensional and multivariate methods can improve statistical power to identify genes with a pleiotropic effect, as illustrated by the identification of *IL9* polymorphisms on both lung function and polysensitization,⁶⁹ and this can be also true for GxE interactions. Although a number of multivariate methods have been proposed, improved methods are needed to integrate information from various data types (including “-omics” data). Combining data from different studies is often needed to get adequate sample size. However, one challenge can be the lack of uniformity in methods used for collecting phenotypic and environmental data across studies. Initiatives such as the Public Population Project in Genomics can greatly facilitate data harmonization and integration.⁷⁰

Regarding the environment, toxicology started as a science of acute effects and was founded on the paradigm that the dose makes the poison: “*Alle Ding sind Gift, und nichts ohn Gift; allein die Dosis macht, das ein Ding kein Gift ist*,” according to Paracelsus,⁷¹ a true European researcher from the 16th century. As shown in Fig 1, the question of the dose has been refined by the consideration of, for example, both the environmental dose and the internal dose. It is therefore not new to consider that continuous assessment of exposure is a key element, and there is a need to go beyond a dichotomous definition of exposed versus unexposed. Biomarkers are not used enough in GxE interaction studies (only in 2 of our 11 sampled studies). Several problems limit their use: the availability of the appropriate markers for the question under study, the availability of the environmental or biological samples to do the measurements, and the fact that they often

TABLE II. Diversity of designs and strategies in the GxE literature illustrated by 11 recently published articles*

Author	Population	Phenotype	Environment	Genes	Hypothesis/ strategy for testing GxE interaction	Authors' conclusions	Issues
Braüner et al, 2011 (Denmark) ⁵⁸	Nested case (1,111)–control (734) study	Wheeze at 18 mo of age	ETS, pets	<i>ORMDL3</i> (1 SNP)	Known marginal effects of G and E	No interaction for ETS but interaction for furred pets	Issue of independence of genes and environment (pets, smoking avoidance)
Bunyavanich et al, 2012 (United States) ⁵⁹	Clinical trial, 422 asthma trios	Lung function	HDM (<i>Dermatophagoides pteronyssinus</i> in µg/g)	<i>P2Y12</i> (19 SNPs from Illumina HumanHap550)	Candidate interaction functional study	Interaction	Role of gene evidenced by considering interaction with environment
Ege et al, 2011 (Germany, Switzerland, Austria) ⁴⁰	Rural children (1,709, including 850 AS+)	Childhood asthma, specific IgE	Farm	Candidate genes and genome- wide SNPs (Illumina Human610 quad array)	GEWIS with a 2-step analysis strategy	No interaction with common genetic variants	Power issue in GEWIS 2-step strategy to reduce multiple testing problem; cautious interpretation
Hwang et al, 2012 (Taiwan) ⁴³	188 AS+, 376 controls	Incident asthma	Mold	Candidate gene <i>IL4</i> (1 SNP)	Marginal effects of G and E	Joint effect of G and E	Scale for interaction
Lee et al, 2012 (Korea) ⁶⁰	9,643 children	Atopic dermatitis	Mold, moving	None (reported parental history)	Indirect	Suggestive interaction	Recall bias (possibly differential)
Liu et al, 2011 (United States, Finland) ⁶¹	387 AS+ 212 AS–; 1,716 AS+, 16,888 AS–	Asthma	Smoking	Candidate gene <i>TSLP</i> (10 SNPs)	Marginal effects	Interaction	Difficulty in replication
Muñoz et al, 2012 (Mexico) ⁶²	90 AS+, 111 AS–	Asthma	ETS (cotinine)	3 genes (4 SNPs)	Candidate genes	No interaction	Power Association ≠ causation
Smit et al, 2012 (8 European countries) ⁶³	844 AS+, 2,046 controls	Cough by asthma status	Smoking, occupational exposures, irritants	3 TRP genes, (SNPs from Illumina Human610 quad array)	Candidate interaction	Suggestive interaction	Difficulty to assess specific exposures
Su et al, 2012 (Taiwan) ⁴⁴	235 AS+, 1,310 AS–	Asthma	9 factors (ETS, pets, household characteristics)	13 candidate genes (17 SNPs)	GxG and GxE with 2- and 3-way interactions	Interaction for dampness	Data reduction approach
Ungvári et al, 2012 (Hungary) ⁶⁴	307 cases, 344 controls	Asthma	Air pollution	<i>NFE2L2</i> , <i>KEAP1</i> , (12 SNPs)	Case-only strategy to test candidate interaction	Interaction (≠ hypothesis)	Time
Zhang et al, 2011 (Russia, Finland) ²⁰	858 children in Karelia	Asthma and allergy	Eastern vs western environment	15 innate immunity candidate genes (24 SNPs)	Pseudoexperimental design in subjects with same ancestry	Interaction	Relevance of macroenvironment

AS+, Subjects with asthma; AS–, subjects without asthma; E, environment; G, gene; GxG, gene-gene interaction; HDM, house dust mite; *KEAP1*, kelch-like ECH-associated protein 1 gene; *NFE2L2*, nuclear factor (erythroid-derived 2)–like 2 gene; *ORMDL3*, ORM1-like 3 (*Saccharomyces cerevisiae*) gene; *P2Y12*, purinergic receptor P2Y, G-protein coupled, 12 gene; *TRP*, transient receptor potential; *TSLP*, thymic stromal lymphopoietin gene.

*The 11 articles shown in this table are all those published in 2011 or 2012 that appeared (September 8, 2012) both in the Hugenet category of GxE interaction for asthma and in PubMed considering “asthma AND gene AND environment AND interaction” key words.

reflect recent exposure only. Indeed, environmental exposures of interest are often chronic exposures, and the window of exposure has emerged as being critical. It is now considered that it might well be the time of exposure instead of the dose that is most crucial (Fig 3). The window of exposure in the subject's life course might influence critical aspects of lung growth,⁷² immunologic maturation,⁷³ epigenetic marks,⁷⁴ and asthma.⁷⁵ More generally, it is critical in the now accepted paradigm of the developmental origin of health and disease.⁷⁶ Longitudinal studies with collection of biological samples is the way forward, but validated

retrospective information obtained by means of questionnaire is also relevant in this context. Clearly, multidisciplinary research is needed to properly assess the environment. When switching from the subject to the population, the issue of a cohort effect regarding exposure emerges. Studies on traditional farming, changing *in utero* exposure to tobacco smoke after cohort changes in female smoking, and changes in industrial processes leading to modification of occupational hazards and biomass or traffic-related pollution exposure, have strong cohort determinants. Tentative solutions include long-term longitudinal studies

TABLE III. Main issues encountered to detect GxE interactions and tentative solutions

Type of issue	Tentative solutions
Phenotype	
Loose phenotype definition	Improve phenotyping through intermediate phenotypes (eg, biological measures of allergy, inflammation, and bronchial responsiveness), subtype of disease defined functionally and pathologically by a molecular mechanism, new phenotypes defined by using clustering analysis (with symptomatic, biological, functional, and therapeutic data), longitudinal data
Time dependence	Time to disease onset as the outcome for analysis; longitudinal studies, repeated measurements
Single phenotype	Joint analysis of multiple phenotypes (eg, combined analysis of lung function data and allergy characteristics) using various techniques (multivariate analysis, univariate analysis of principal components of the original phenotypes followed by combination of the independent test statistics associated with the principal components, combination of test statistics from univariate analysis of each phenotype taking into account the correlation among the test statistics)
Heterogeneity in data collection	Data harmonization, use of standardized procedures
Environment	
Loose definition of exposure and measurement error	Use of continuous or at least categorical variables to assess exposure (notion of dose), more investment in assessing environment, use of biomarkers
Time dependence (window of exposure)	Repeated measures of environment Longitudinal studies
Cohort effect regarding exposure	Studies incorporating macroenvironment (broad lifestyle, studies of migrants) or considering subjects of various ages or using retrospective information
Modifying effect by another environmental factor	Assess relevant exposures, multilevel modeling (eg, using country of residence and familial and individual characteristics), give more consideration to multiple exposures (eg, air pollution, smoking, and diet)
Social determinants of exposures at the collective and individual levels	Consider social determinants, possibly multilevel modeling (eg, area of residence and individual levels), use of causal models, stratified analysis
Selection bias: healthy smoker effect or work avoidance	Correction by causal modeling (attempting to correct the observations as if the bias had not occurred)
Gene	
Loose definition of genetic variant: SNPs are proxies of causal variants	Increase SNP density by imputing nongenotyped SNPs using genotyped SNPs and reference panels (eg, HapMap2) Sequencing toward characterization of causal variants Combining GWAS of disease and GWAS of gene expression
Single-SNP analysis	Joint analysis of multiple SNPs; pathway-based approaches (analyses of sets of SNPs [or genes] belonging to the same biological pathway) incorporating GxE interaction
Gene-gene interaction	Complex modeling, building interrelationships among genes and gene expression profiles (gene networks)
Structural variation (CNVs)	Improve methods to infer CNVs from SNP arrays
Rare variants	Improve quality control of data generated by using new sequencing technologies, use appropriate study design and development of statistical methods to increase power
Gene function unknown	Gene expression studies, functional studies with animal models
GxE interaction	
Power	Defining optimal study designs (eg, case only, case-control, cohort designs) Development of statistical methods to increase power Meta-analyses of existing studies for GxE interaction studies
Multiplicity of statistical methods	Comparison of methods through simulations Combination of strategies and methods
Use of biological knowledge	Pathway-based approaches, hierarchical modeling, data-reduction methods (to restrict investigation of GxE to a subset of relevant genes and/or environmental factors) Improve gene annotations and other sources of biological information

and studies of subjects of various ages, including retrospective information on the childhood environment, interest in the macroenvironment,²⁰ and international studies, such as the International Study of Asthma and Allergies in Childhood.⁷⁷

The environmental factor under study in a given GxE interaction analysis might not be the sole exposure to which the subjects are subjected, but other exposures need to be taken into account in the design, analysis, and interpretation of the results. One clear

situation is when the effect of an environmental factor might be prevented by another one, such as oxidant exposure caused by air pollution and consumption of fruits with antioxidant properties.⁷⁸

A pathway-based approach, in which several genes and environmental factors related to that biological pathway are studied together, is one way to go forward. Even with such approaches, difficulties can remain because protective factors counterbalancing deleterious factors are often unknown or data are unavailable.

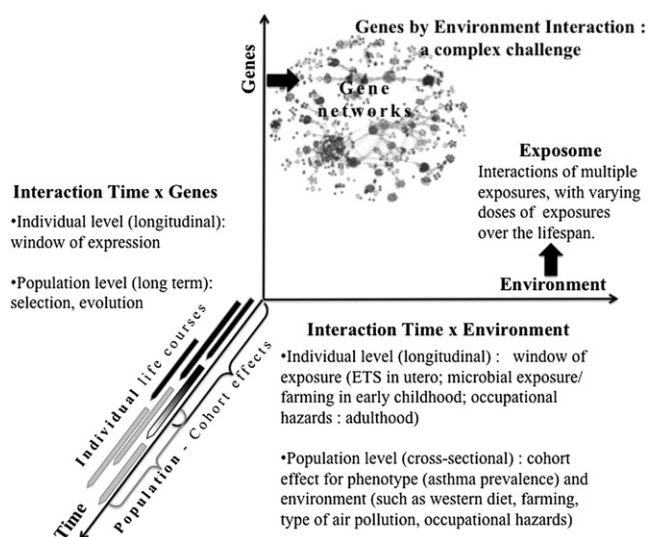


FIG 3. GxE interactions: a complex challenge including time. The time axis represents the individual life course, with changes in environment corresponding to cohort effects (black and light gray lines). Within their life course, some subjects can experience cohort-like acute changes in case of migration or cohort-like subacute changes in epidemiologic transitional situations occurring in developing countries. Such changes can result in new combinations of exposures at a given time or over the life course. Time is of particular importance in the context of the Developmental Origins of Health (DoHAD) hypothesis.⁷⁶ Disappearance of protective factors (eg, modification of the microbiome) and emergence of new deleterious factors (eg, new occupational cleaning agents) are cohort related and influence the windows of expression of asthma. Gene effect on disease can also vary with time, either at the individual level (regulation of gene expression) or at the population level (selection and evolution). The concept of the exposome⁴² aims at integrating multiple exposures over the lifespan. Finally, phenotypes will result in the interactions of gene networks, in which homeostasis will have a regulatory effect, with potential infinite interaction events with various elements of the exposome. Although ambitious approaches to address various aspects simultaneously are emerging, the challenge remains formidable and, as described by Descartes,⁸ it is necessary to combine realistic research strategies to disentangle the puzzle.

Spatial clustering of air pollution, temperature, and social determinants can lead to confounding effects. Air pollution is assessed through residential information. Determination of interactions of air pollution with genetic factors might be difficult for assessment of whether asthma in a given geographic area depends more on social determinants or their consequences (eg, obesity) than on air pollution. Multilevel modeling (spatial clustering) is one possibility to overcome this issue. Examination of numerous environmental factors simultaneously might be feasible in future studies, and this is illustrated by ongoing projects that are focusing on the microbiome.⁷⁹ Integration of various amounts of available knowledge or totally agnostic approaches on the environmental side can be envisioned, but it is clear that the level of exhaustivity on the environmental side would never reach the completeness of DNA variation. However, there have been, until now, too few studies jointly examining a large number of environmental factors, even when they are available.

It is worth noting that methods defining clustering patterns or data-reduction approaches are currently used in some areas but not in others, likely because of habits rather than rationale. It is common to define dietary patterns by using a variety of multivariate methods.⁸⁰ There is an increasing use of cluster analysis applied to multiple phenotypes and principal components

analysis applied to genome-wide SNP data to take into account population structure, but similar methods need to be more broadly applied to environmental exposures. Of specific interest are the methods that can incorporate previous knowledge. Social determinants of environmental exposures can occur at the collective level (regulations regarding air pollution or occupational exposures, price of cigarettes, smoking-free areas, biomass fuel exposure, food, or transport availability) and at the individual level (diet, obesity, changing work, stopping smoking, and moving). It is therefore important to consider these social determinants, possibly through stratified analyses, multilevel modeling, and increasing use of causal models.

Other peculiarities related to environmental factors pose problems in the design and analysis of GxE interactions. For example, the case-only design or a 2-step approach proposed in the GEWIS model to test for GxE interaction implies the independence of genes and environment in the population.¹⁵ However, social determinants might lead to nonindependence between genes and environment in the population, as is the case for smoking dependence. Smoking habits are influenced by both genes and social determinants at the individual and population levels. This complicates the analysis of GxE interactions, but use of appropriate study designs and analytic methods (eg, adjustment and causal modeling) can overcome this issue. Similarly, selection biases can occur (healthy worker effect and moving in relation to air pollution annoyance). In family-based studies the classical aspect of the shared environment within families should be handled with direct assessment of that environment. Regarding lifestyle, consideration should also be given to assortative mating, which is well known to occur in relation to risk factors of asthma, such as smoking or obesity.⁸¹

Up to now, studies of GxE interactions have examined single genetic markers one at a time within candidate genes or at the genome-wide level (GEWIS). In candidate gene studies candidate polymorphisms have most often been chosen because of their potential functional role in disease, whereas SNPs in genome-wide studies are most often proxies (or tags) associated with unknown functional variants. This can be considered a type of a loose definition of a genetic marker (or measurement error), as for an ill-defined phenotype or environmental exposure. Loci detected by using GWASs are often quite broad, spanning hundreds of kilobases and including a number of potential candidate genes. Efforts to fine map the candidate causal variants responsible for GWAS association signals have had limited success up to now. However, the joint analysis of multiple genetic markers rather than single-marker analysis within a locus evidenced by a GWAS has recently proved to be a promising approach to pinpoint a gene with a potential functional role in disease (Brossard et al, personal communication). Combining data from GWASs of both disease and gene expression levels provides another route to shed light on the potential functional role of genes within GWAS loci, as illustrated by the identification of ORM1-like 3 (*Saccharomyces cerevisiae*) gene (*ORMDL3*) variants that increase the risk of asthma and are associated with *ORMDL3* gene expression.^{46,82} In the post-GWAS era, efforts are now focused on developing multimarker methods to mine the mass of association signals that do not reach genome-wide significance but, if considered jointly, might indicate the involvement of relevant sets of genes or pathways in disease. These methods differ according to the study design, analysis strategy (1-step vs 2-step approaches), use (or not use) of biological knowledge, and the statistical methodology used.^{83,84} Incorporating GxE interactions into pathway-based

analysis can provide a promising alternative between candidate interaction studies and GEWISs, as illustrated by the investigation of interactions of genes in the oxidative-stress pathway with particulate matter and tobacco smoke exposure in lung function decrease.⁸⁵

Genes can also interact with other genes, which adds another level of complexity. Models and methods to address epistasis have been extensively reviewed.⁸⁶ In addition to SNPs, other types of genetic variation can be involved, such as copy number variants (CNVs), which can also interact with environmental factors. However, methods used to determine CNV location and copy number from SNP array data have to be improved to obtain high-performance CNV analysis.⁸⁷ Recent advances in sequencing technologies have invigorated the search for rare variants in the cause of complex diseases.⁸⁸ This raises a number of issues in terms of defining quality control criteria for sequenced data, study designs, and statistical methodology to obtain adequate power. There is currently active research to address these problems.¹⁴

One of the major issues raised by studies of GxE interaction is statistical power. Much larger sample sizes are needed to detect interactions than main effects: an interaction requires at least a 4-fold larger sample size than does a main effect of comparable magnitude (eg, thousands of cases and control subjects to detect an interaction relative risk of 1.5 for a candidate study and tens of thousands for a GEWIS). Advantages and disadvantages of various study designs for detecting GxE interactions have been discussed in previous review articles.^{13-15,56} Increased power can be achieved by enriching the sample for subjects with a family history of disease or one having extreme phenotypes or extreme environmental exposures. Many statistical methods have recently been proposed to enhance power of detecting GxE interactions in GEWISs.⁸⁹⁻⁹³ Two-step strategies have been proposed^{94,95} and used in the farming exposure–asthma GEWI to reduce the problem of multiple testing that is crucial in GEWISs.⁴⁰ The power of these approaches has been extensively studied through simulations that have shown that optimal strategies depend on the underlying model.⁹⁶ Because the true interaction models are unknown, “cocktail” methods that combine each method’s most appealing aspect can provide a comprehensive and powerful tool for genome-wide detection of GxE interactions.⁹⁷ Moreover, as for GWASs,^{98,99} meta-analysis of GEWISs can increase power further and provide robust estimates of GxE interaction.^{100,101}

As an alternative to purely agnostic approaches, pathway-based approaches,^{83-85,102} data-reduction methods,⁴⁵ or hierarchical modeling¹⁰³ can be used advantageously for detecting GxE interactions. All pathway-based methods rely on biological knowledge that is stored in databases, such as Gene Ontology, Kyoto Encyclopedia of Genes and Genomes, and Biocarta. These databases often have inconsistencies in gene nomenclature, and a proportion of known genes do not map to functional annotations within these databases. Therefore improved annotation is needed, along with other sources of biological information, to better assign pathways. Finally, experimental studies¹⁰⁴ and functional studies^{17,37} are needed to validate GxE interactions.

HOW TO INTEGRATE NEW KNOWLEDGE AND NEW TOOLS?

Advances in high-throughput sequencing technologies and the availability of a massive amount of genome-scale data at various levels (DNA [genomics], RNA [transcriptomics], and protein

[proteomics]) provide new opportunities for disentangling the complex mechanisms underlying complex diseases, such as asthma, and for progressing rapidly in medical and public health applications (the so-called personalized medicine). Strategies integrating various “-omics” data have been emerging¹⁰⁵ but need to be further developed. The understanding of the mechanisms underlying the interplay between genes and the environment can be provided by epigenomics.¹⁰⁶ It is more and more believed that epigenetic mechanisms (ie, functionally relevant modifications to the genome that do not involve a change in the nucleotide sequence) play a major role in gene regulation and are implicated in the cause of many diseases, including asthma and allergy. Epigenetics is known to play a role in T-cell differentiation and regulation,¹⁰⁷ and environmental factors associated with asthma, such as exposure to tobacco smoke, ozone, oxidative stress, endotoxin, and/or diet, have been reported to induce epigenetic changes in experimental settings.¹⁰⁸ Observations, such as the maternal effect of asthma and genetic imprinting at specific loci, are also consistent with epigenetic effects in asthmatic patients.^{109,110} New technologies allow us to explore epigenetic alterations, such as patterns of DNA methylation across the genome. This makes it possible not only to identify the genetic and environmental components of DNA methylation patterns but to understand how these methylation patterns mediate the effect of known environmental factors and GxE interactions on asthma-related phenotypes.⁷⁴ There is also accumulating evidence that the gut microbiome influences the development of inflammatory disease,⁷⁹ and the composition of the gut microbiome, the airway microbiome, or both might be important in determining the risk of asthma development and could contribute to chronic asthma.¹¹¹ The gut and airway microbiota initially acquired during infancy are likely to be modified by environmental exposures, such as exposure to pet dogs or farm animals, which are known to protect against asthma.¹¹² Thus new sequencing technology that enables us to study the microbiome on a large scale opens up new routes of research into the understanding of GxE interactions. One aspect that needs consideration is the specificity of tissues in which these molecular mechanisms can occur and make the acquisition of samples of various sorts particularly challenging.

New approaches need to be developed to integrate these various molecular data and to link them to the disease process. Data-mining methods and methodology for building molecular networks can be instrumental in this endeavor.¹¹³ A comprehensive understanding of the interactions among the genome, the proteome, the environment, and multiple disease-related phenotypes (or pathophenome) mediated by the underlying molecular and cellular network is more and more seen as the way forward to achieve insight into the complex mechanisms underlying the disease process and to progress toward personalized intervention. As stated by Barabasi et al,¹¹⁴ the “think globally, act locally” paradigm of network medicine can allow identification of new drug targets and biomarkers for complex diseases and to generate local interventions that might cure the disease.

CONCLUSION

Although the global vision of “-omics” techniques and a molecular network-based approach to human disease are very appealing for evolving toward personalized medicine, numerous pieces of the whole system are still missing and most likely need to

be tackled individually. Resistance to approaches only driven by technology is warranted, while keeping an open mind on approaching the various parts of the puzzle with methodologic rigor, as described by Descartes.⁴ International and multidisciplinary approaches should be developed without the illusion of a rapid comprehensive answer. Results regarding genetic associations, GxE interactions, or both should be interpreted with a healthy respect for the limitations of the methods and “without replacing irrational exuberance by imprudent pessimism” (as stated by Herrington,¹¹⁵ even before the GWAS era). The future of a molecular network-based approach of disease to which GxE interactions are related requires space for innovative research. Assembling the various parts of a puzzle in a complex system might well occur in a way that might not necessarily follow the rules of logic.

What do we know?

- Epidemiologic studies have evidenced many environmental exposures associated with asthma.
- Genetic studies have identified a number of genetic loci associated with asthma.
- Windows of exposure and windows of expression of the disease over the lifespan are critical issues to be considered.
- Few GxE interactions have been established.

What is still unknown?

- Joint effects of environmental factors, genetic factors, or both on asthma have been studied rarely.
- Interactions among genes, environmental factors, and GxE interactions are poorly documented.

REFERENCES

- Haldane JBS. Heredity and politics. London: George Allen and Unwin Ltd; 1938.
- Khoury MJ. Relationship between medical genetics and public health: changing the paradigm of disease prevention and the definition of genetic disease. *Am J Med Genet* 1997;71:289-91.
- Shostak S. Locating gene-environment interaction: at the intersections of genetics and public health. *Soc Sci Med* 2003;56:2327-42.
- Descartes R. Discours de la méthode. Pour bien conduire la raison et chercher la vérité dans les sciences, La Haye, 1637. Available at: <http://gallica.bnf.fr>. Accessed August 30, 2012. English translation: Discourse on the method on rightly conducting the reason and seeking the truth in the sciences. Available at: <http://www.gutenberg.org/ebooks/59>. Accessed August 30, 2012.
- Schork NJ. Genetics of complex disease: approaches, problems and solutions. *Am J Respir Crit Care Med* 1997;156(suppl):S103-9.
- Anderson GP. Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. *Lancet* 2008;372:1107-19.
- Schulte PA. A conceptual framework for the validation and use of biologic markers. *Environ Res* 1989;48:129-44.
- Schulte PA, Lomax GP, Ward EM, Colligan MJ. Ethical issues in the use of genetic markers in occupational epidemiologic research. *J Occup Environ Med* 1999;41:639-46.
- Sharp RR. Ethical issues in environmental health research. *Environ Health Perspect* 2003;111:1786-8.
- Khoury MJ, Gwinn M, Clyne M, Yu W. Genetic epidemiology with a capital E, ten years after. *Genet Epidemiol* 2011;35:845-52.
- Thomas DC. Genetic epidemiology with a capital E: where will we be in another 10 years? *Genet Epidemiol* 2012;36:179-82.
- Gene, environment and health initiative; 2006. Available at: <http://www.genome.gov/19518663>. Accessed August 30, 2012.
- Bookman EB, McAllister K, Gillanders E, Wanke K, Balshaw D, Rutter J, et al. Gene501 environment interplay in common complex diseases: forging an integrative model—recommendations from an NIH workshop. *Genet Epidemiol* 2011;35:217-25.
- Mechanic LE, Chen HS, Amos CI, Chatterjee N, Cox NJ, Divi RL, et al. Next generation analytic tools for large scale genetic epidemiology studies of complex diseases. *Genet Epidemiol* 2012;36:22-35.
- Thomas D. Gene-environment-wide association studies: emerging approaches. *Nat Rev Genet* 2010;11:259-72.
- Ottman R. An epidemiologic approach to gene-environment interaction. *Genet Epidemiol* 1990;7:177-85.
- Ober C, Vercelli D. Gene-environment interactions in human disease: nuisance or opportunity? *Trends Genet* 2011;27:107-15.
- Masoli M, Fabian D, Holt S, Beasley R. The global burden of asthma: executive summary of the GINA Dissemination Committee Report. *Allergy* 2004;59:469-78.
- Candelaria PV, Backer V, Khoo SK, Bizzintino JA, Hayden CM, Baynam G, et al. The importance of environment on respiratory genotype/phenotype relationships in the Inuit. *Allergy* 2010;65:229-37.
- Zhang G, Candelaria P, Mäkelä JM, Khoo S-K, Hayden MC, von Hertzen L, et al. Disparity of innate immunity-related gene effects on asthma and allergy on Karelia. *Pediatr Allergy Immunol* 2011;22:621-30.
- Weiss ST. New approaches to personalized medicine for asthma: Where are we? *J Allergy Clin Immunol* 2012;129:327-34.
- Kauffmann F, Nadif F. Candidate gene-environment interactions. *J Epidemiol Community Health* 2010;64:188-9.
- Vercelli D. Learning from discrepancies: *CD14* polymorphisms, atopy and the endotoxin switch. *Clin Exp Allergy* 2003;33:153-5.
- Martinez FD. *CD14*, endotoxin, and asthma risk: actions and interactions. *Proc Am Thorac Soc* 2007;4:221-5.
- Howell WM, Holgate ST. *HLA* genetics and allergic disease. *Thorax* 1995;50:815-8.
- Aderem A, Ulevitch RJ. Toll-like receptors in the induction of the innate immune response. *Nature* 2000;406:782-7.
- Bouzigon E, Corda E, Aschard H, Dizier MH, Boland A, Bousquet J, et al. Effect of 17q21 variants and smoking exposure in early-onset asthma. *N Engl J Med* 2008;359:1985-94.
- Koppelman GH. Gene-environment interaction in allergic disease: more questions, more answers. *J Allergy Clin Immunol* 2007;120:1266-8.
- Martinez FD. Genes, environments, development and asthma: a reappraisal. *Eur Respir J* 2007;29:179-84.
- Castro-Giner F, Künzli N, Jacquemin B, Forsberg B, de Cid R, Sunyer J, et al. Traffic-related air pollution, oxidative stress genes, and asthma (ECHRS). *Environ Health Perspect* 2009;117:1919-24.
- Le Souëf PN. Gene-environmental interaction in the development of atopic asthma: new developments. *Curr Opin Allergy Clin Immunol* 2009;9:123-7.
- London S, Romieu I. Gene by environment interaction in asthma. *Annu Rev Public Health* 2009;30:55-80.
- Schwartz DA. Gene-environment interactions and airway disease in children. *Pediatrics* 2009;123(suppl):S151-9.
- Von Mutius E. Gene-environment interactions in asthma. *J Allergy Clin Immunol* 2009;123:3-11.
- Kauffmann F, Castro-Giner F, Smit LAM, Nadif R, Kogevinas M. Gene-environment interaction in occupational asthma. In: Sigsgaard T, Heederik D, editors. *Occupational asthma*. Basel: Birkhäuser/Springer Verlag AG; 2010. p. 205-28.
- Romieu I, Moreno-Macias H, London SJ. Gene by environment interaction and ambient air pollution. *Proc Am Thorac Soc* 2010;7:116-22.
- Vercelli D. Gene-environment interactions in asthma and allergy: the end of the beginning. *Curr Opin Allergy Clin Immunol* 2010;10:145-8.
- Kim SH, Cho BY, Park CS, Shin ES, Cho EY, Yang EM, et al. Alpha-T-catenin (*CTNNA3*) gene was identified as a risk variant for toluene diisocyanate-induced asthma by genome-wide association analysis. *Clin Exp Allergy* 2009;39:203-12.
- Kim JH, Park BL, Cheong HS, Bae JS, Park JS, Jang AS, et al. Genome-wide and follow-up studies identify *CEP68* gene variants associated with risk of aspirin-intolerant asthma. *PLoS One* 2010;5:e13818.
- Ege MJ, Strachan DP, Cookson WO, Moffatt MF, Gut I, Lathrop M, et al. Gene-environment interaction for childhood asthma and exposure to farming in Central Europe. *J Allergy Clin Immunol* 2011;127:138-44.
- Ramasamy A, Kuokkanen M, Vedantam S, Gajdos ZK, Couto Alves A, Lyon HN, et al. Genome wide association studies of asthma in population-based cohorts confirm known and suggested loci and identify an additional association near *HLA*. *PLoS One* 2012;7:e44008.
- Wild CP. The exposome: from concept to utility. *Int J Epidemiol* 2012;41:24-32.
- Hwang BF, Liu IP, Huang TP. Gene-environment interaction between interleukin-4 promoter and molds in childhood asthma. *Ann Epidemiol* 2012;22:250-6.

44. Su MW, Tung KY, Liang PH, Tsai CH, Kuo NW, Lee YL. Gene-gene and gene-environment interactions of childhood asthma: a multifactor dimension reduction approach. *PLoS One* 2012;7:e30694.
45. Ritchie MD, Moutsier AA. Multifactor dimensionality reduction for detecting gene-gene and gene-environment interaction in pharmacogenomics studies. *Pharmacogenomics* 2005;6:823-34.
46. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating *ORMDL3* expression contribute to the risk of childhood asthma. *Nature* 2007;448:470-3.
47. Halapi E, Gudbjartsson DF, Jostdottir GM, Bjornsdottir US. 17q12-21 variants interact with smoke exposure as a risk factor for pediatric asthma but are equally associated with early-onset versus late-onset asthma in North Americans of European ancestry. *J Allergy Clin Immunol* 2009;124:605-7.
48. van der Valk RJP, Duijts L, Kerkhof M, Willemsen SP, Hofman A, Moll HA, et al. Interactions of a 17q12 variant with both fetal and infant smoke exposure in the development of childhood asthma-like symptoms. *Allergy* 2012;67:767-74.
49. Colilla S, Nicolae D, Pluzhnikov A, Blumenthal MN, Beaty TH, Bleecker ER, et al. Evidence for gene-environment interactions in a linkage study of asthma and smoking exposure. *J Allergy Clin Immunol* 2003;111:840-6.
50. Meyers DA, Postma DS, Stine OC, Koppelman GH, Ampleford EJ, Jongepier H, et al. Genome screen for asthma and bronchial hyperresponsiveness: interactions with passive smoke exposure. *J Allergy Clin Immunol* 2005;115:1169-75.
51. Dizier MH, Bouzigon E, Guilleud-Bataille M, Siroux V, Lemaître A, Boland A, et al. Evidence for gene x smoking exposure interactions in a genome-wide linkage screen of asthma and bronchial hyper-responsiveness in EGEA families. *Eur J Hum Genet* 2007;15:810-5.
52. Khoury MJ, Wacholder S. From genome-wide association studies to gene environment-wide interaction studies—challenges and opportunities. *Am J Epidemiol* 2009;169:227-30.
53. Thomas D. Methods for investigating gene-environment interactions in candidate pathway and genome-wide association studies. *Annu Rev Public Health* 2010;31:21-36.
54. Hamza TH, Chen H, Hill-Burns EM, Rhodes SL, Montimurro J, Kay DM, et al. Genome-wide gene-environment study identifies glutamate receptor gene *GRIN2A* as a Parkinson's disease modifier gene via interaction with coffee. *PLoS Genet* 2011;7:e1002237.
55. Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med* 2012;18:716-25.
56. Aschard H, Lutz S, Maus B, Duell EJ, Fingerlin TE, Chatterjee N, et al. Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. *Hum Genet* 2012;131:1591-613.
57. Yu W, Gwinn M, Clyne M, Yesupriya A, Khoury KJ. A Navigator for Human Genome Epidemiology. *Nat Genet* 2008;40:124-5.
58. Bräuner EV, Loft S, Raaschou-Nielsen O, Vogel U, Andersen PS, Sørensen M. Effects of a 17q21 chromosome gene variant, tobacco smoke and furred pets on infant wheeze. *Genes Immun* 2012;13:94-7.
59. Bunyavanich S, Boyce JA, Raby BA, Weiss ST. Gene-by-environment effect of house dust mite on purinergic receptor P2Y12 (P2RY12) and lung function in children with asthma. *Clin Exp Allergy* 2012;42:229-37.
60. Lee JY, Seo JH, Kwon JW, Yu J, Kim BJ, Lee SY, et al. Exposure to gene-environment interactions before 1 year of age may favor the development of atopic dermatitis. *Int Arch Allergy Immunol* 2012;157:363-71.
61. Liu M, Rogers L, Cheng Q, Shao Y, Fernandez-Beros ME, Hirschhorn JN, et al. Genetic variants of *TSLP* and asthma in an admixed urban population. *PLoS One* 2011;6:e25099.
62. Muñoz B, Magaña JJ, Romero-Toledo I, Juárez-Pérez E, López-Moya A, Leyva-García N, et al. The relationship among *IL-13*, *GSTP1*, and *CYP1A1* polymorphisms and environmental tobacco smoke in a population of children with asthma in Northern Mexico. *Environ Toxicol Pharmacol* 2012;33:226-32.
63. Smit LA, Kogevinas M, Antó JM, Bouzigon E, González JR, Le Moual N, et al. Transient receptor potential genes, smoking, occupational exposures and cough in adults. *Respir Res* 2012;13:26.
64. Ungvári I, Hadadi E, Virág V, Nagy A, Kiss A, Kalmár A, et al. Relationship between air pollution, *NFE2L2* gene polymorphisms and childhood asthma in a Hungarian population. *J Community Genet* 2012;3:25-33.
65. Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, et al. Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* 2008;178:218-24.
66. Moore WC, Meyers DA, Wenzel SE, Teague WG, Li H, Li X, et al. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* 2010;181:315-23.
67. Siroux V, Basagaña X, Boudier A, Pin I, Garcia-Aymerich J, Vesin A, et al. Identifying adult asthma phenotypes using a clustering approach. *Eur Respir J* 2011;38:310-7.
68. Forno E, Lasky-Su J, Himes B, Howrylak J, Ramsey C, Brehm J, et al. Genome-wide association study of the age of onset of childhood asthma. *J Allergy Clin Immunol* 2012;130:83-90.
69. Aschard H, Bouzigon E, Corda E, Ulgen A, Dizier MH, Gormand F, et al. Sex-specific effect of *IL9* polymorphisms on lung function and polysensitization. *Genes Immun* 2009;10:559-65.
70. Knoppers BM, Fortier I, Legault D, Burton P. The Public Population Project in Genomics (P3G): a proof of concept. *Eur J Hum Genet* 2008;16:664-5.
71. Borzelleca JF. Paracelsus: herald of modern toxicology. *Toxicol Sci* 2000;53:2-4.
72. Becklake MR, Kauffmann F. Gender differences in airway behaviour over the human life span. *Thorax* 1999;54:1119-38.
73. Sly PD, Holt PG. Role of innate immunity in the development of allergy and asthma. *Curr Opin Allergy Clin Immunol* 2011;11:127-31.
74. Yang IV, Schwartz DA. Epigenetic mechanisms and the development of asthma. *J Allergy Clin Immunol* 2012 [Epub ahead of print].
75. Neuman A, Hohmann C, Orsini N, Pershagen G, Eller E, Fomsgaard Kjaer H, et al. Maternal smoking in pregnancy and asthma in preschool children: a pooled analysis of 8 birth cohorts. *Am J Respir Crit Care Med* 2012 [Epub ahead of print].
76. Barouki R, Gluckman PD, Grandjean P, Hanson M, Heindel JJ. Developmental origins of non communicable disease: implications for research and public health. *Environ Health* 2012;11:42.
77. Lewis S. ISAAC—a hypothesis generator for asthma? *Lancet* 1998;351:1220-1.
78. Romieu I, Sienra-Monge JJ, Ramírez-Aguilar M, Moreno-Macías H, Reyes-Ruiz NI, Estela del Río-Navarro B, et al. Genetic polymorphism of *GSTM1* and antioxidant supplementation influence lung function in relation to ozone exposure in asthmatic children in Mexico City. *Thorax* 2004;59:8-10.
79. Renz H, von Mutius E, Brandtzaeg P, Cookson WO, Autenrieth IB, Haller D. Gene-environment interactions in chronic inflammatory disease. *Nat Immunol* 2011;12:273-7.
80. Varraso R, Garcia-Aymerich J, Monier F, Le Moual N, De Batlle J, Miranda G, et al. Assessment of dietary patterns in nutritional epidemiology: principal component analysis versus confirmatory factor analysis. *Am J Clin Nutr* 2012;96:1079-92.
81. Sutton GC. Assortative marriage for smoking habits. *Ann Hum Biol* 1980;7:449-56.
82. Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, Wong KC, et al. A genome-wide association study of global gene expression. *Nat Genet* 2007;39:1202-7.
83. Cantor RM, Lange K, Sinheimer JS. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *Am J Hum Genet* 2010;86:6-22.
84. Holmans P. Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits. *Adv Genet* 2010;72:141-79.
85. Curjuric I, Imboden M, Nadif R, Kumar A, Schindler C, Haun M, et al. Different genes interact with particulate matter and tobacco smoke exposure in affecting lung function decline in the general population. *PLoS One* 2012;7:e40175.
86. Cordell H. Detecting gene-gene interactions that underlie human diseases. *Nat Rev Genet* 2009;10:392-404.
87. Marenne G, Rodríguez-Santiago B, Closas MG, Pérez-Jurado L, Rothman N, Rico D, et al. Assessment of copy number variation using the Illumina Infinium 1M SNP-array: a comparison of methodological approaches in the Spanish Bladder Cancer/EPICURO study. *Hum Mutat* 2011;32:240-8.
88. Yang IV, Schwartz DA. The next generation of complex lung genetic studies. *Am J Respir Crit Care Med* 2012 [Epub ahead of print].
89. Mukherjee B, Ahn J, Gruber SB, Chatterjee N. Testing gene-environment interaction in large-scale case-control association studies: possible choices and comparisons. *Am J Epidemiol* 2012;175:177-90.
90. Li D, Conti L. Detecting gene-environment interactions using a combined case-only and case control approach. *Am J Epidemiol* 2008;169:497-504.
91. Mukherjee B, Chatterjee N. Exploiting gene-environment independence for analysis of case control studies: an empirical Bayes type shrinkage estimator to trade-off between bias and efficiency. *Biometrics* 2008;64:685-94.
92. Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect genetic associations. *Hum Hered* 2007;63:111-9.
93. Roeder K, Wasserman L. Genome-wide significance levels and weighted hypothesis testing. *Stat Sci* 2009;24:398-413.
94. Kooperberg C, LeBlanc M. Increasing the power of identifying gene-gene interactions in genome-wide association studies. *Genet Epidemiol* 2008;32:255-63.
95. Murcray CE, Lewinger JP, Conti DV, Thomas DC, Gauderman WJ. Sample size requirements to detect gene-environment interactions in genome-wide association studies. *Genet Epidemiol* 2011;35:201-10.
96. Dai JY, Logsdon BA, Huang Y, Hsu L, Reiner AP, Prentice RL, et al. Simultaneously testing for marginal genetic association and gene-environment interaction. *Am J Epidemiol* 2012;176:164-73.

97. Hsu L, Jiao S, Dai JY, Hutter C, Peters U, Kooperberg C. Powerful cocktail methods for detecting genome-wide gene-environment interaction. *Genet Epidemiol* 2012;36:183-94.
98. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010;363:1211-21.
99. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011;43:887-92.
100. Manning AK, LaValley M, Liu CT, Rice K, An P, Liu Y, et al. Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP $\times$ environment regression coefficients. *Genet Epidemiol* 2011;35:11-8.
101. Aschard H, Hancock DB, London SJ, Kraft P. Genome-wide meta-analysis of joint tests for genetic and gene-environment interaction effects. *Hum Hered* 2010;70:292-300.
102. Ramanan VK, Shen L, Moore JH, Saykin AJ. Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends Genet* 2012;28:323-32.
103. Hung RJ, Brennan P, Malaveille C, Porru S, Donato F, Boffetta P, et al. Using hierarchical modeling in genetic association studies with multiple markers: application to a case-control study of bladder cancer. *Cancer Epidemiol Biomarkers Prev* 2004;13:1013-21.
104. Willis-Owen SAG, Valdar W. Deciphering gene-environment interactions through mouse models of allergic asthma. *J Allergy Clin Immunol* 2009;123:14-23.
105. Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature* 2009;461:218-23.
106. Rodenhiser D, Mann M. Epigenetics and human disease: translating basic biology into clinical applications. *CMAJ* 2006;174:341-8.
107. Zhu J, Paul WE. CD4 T cells: fates, functions, and faults. *Blood* 2008;112:1557-69.
108. Ho SM. Environmental epigenetics of asthma: an update. *J Allergy Clin Immunol* 2010;126:453-65.
109. Kabesch M, Michel S, Tost J. Epigenetic mechanisms and the relationship to childhood asthma. *Eur Respir J* 2010;36:950-61.
110. Koppelman GH, Nawijn MC. Recent advances in the epigenetics and genomics of asthma. *Curr Opin Allergy Clin Immunol* 2011;11:414-9.
111. Han MK, Huang YJ, Lipuma JJ, Boushey HA, Boucher RC, Cookson WO, et al. Significance of the microbiome in obstructive lung disease. *Thorax* 2012;67:456-63.
112. Ege MJ, Mayer M, Normand AC, Genuneit J, Cookson WO, Braun-Fahrlander C, et al. Exposure to environmental microorganisms and childhood asthma. *N Engl J Med* 2011;364:701-9.
113. Ghazalpour A, Doss S, Zhang B, Wang S, Plaisier C, Castellanos R, et al. Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genet* 2006;2:e130.
114. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* 2011;12:56-68.
115. Herrington D. Eliminating the improbable. Sherlock Holmes and the standards of evidence in the genomic age. *Circulation* 2005;112:2081-4.