Current Challenges and New Opportunities for Gene-Environment Interaction Studies of Complex Diseases

Kimberly McAllister, Leah E. Mechanic, Christopher Amos, Hugues Aschard, Ian A. Blair, Nilanjana Chatterjee, David Conti, W. James Gauderman, Li Hsu, Carolyn M. Hutter, Marta M. Jankowska, Jacqueline Kerr, Peter Kraft, Stephen B. Montgomery, Bhramar Mukherjee, George J. Papanicolaou, Chirag J. Patel, Marylyn D. Ritchie, Beate R. Ritz, Duncan C. Thomas, Peng Wei, John S. Witte on behalf of GxE meeting participants

Running Title: Gene-Environment Interaction Studies of Complex Diseases

Corresponding Author:
Leah E. Mechanic, Ph.D., M.P.H.
Genomic Epidemiology Branch
Epidemiology and Genomics Research Program
Division of Cancer Control and Population Sciences
National Cancer Institute
9609 Medical Center Drive, Rm. 4E104, MSC 9763
Bethesda, MD 20892
(For express delivery, use Rockville, MD 20850)
Phone: 240-276-6847
Email: mechanil@mail.nih.gov

Abbreviations:
CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology; ENCODE, Encyclopedia of DNA Elements; GxE, gene-environment; GWAS, genome-wide association study; GTEx, Genotype-Tissue Expression; TCGA, the Cancer Genome Atlas

Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2017. This work is written by (a) US Government employee(s) and is in the public domain in the US.
Abstract

Recently, many new approaches, study designs, and statistical and analytical methods have emerged for studying gene-environment interactions (GxE) in large-scale human population studies. There are currently opportunities in this field, particularly with respect to the incorporation of -omics and next-generation sequencing data and continual improvement in measures of environmental exposures implicated in complex disease outcomes. A workshop held on October 17-18, 2014 by the National Institute of Environmental Health Sciences and the National Cancer Institute in conjunction with the annual American Society of Human Genetics meeting explored new approaches and tools developed in recent years for GxE interaction discovery. This paper will highlight current and critical issues and themes in GxE research discussed that need additional consideration including the topics of improved data analytical methods, environmental exposure assessment, and incorporation of functional data and annotations.

Keywords: gene-environment, genome-wide association study, environmental exposure
Introduction

Genetic and environmental factors are thought to contribute to the etiology of most complex diseases. Through genome-wide association studies (GWAS), thousands of common loci associated with complex diseases have been identified (1-3). Researchers have been motivated to discover and describe how the interplay of these factors influence disease risk and outcomes. Several reasons for studying gene-environment (GxE) interaction include: providing insights into the biology of disease (e.g. developing new models for disease etiology based on observed GxE findings); building better prognostic models (e.g. using genotype to inform treatment and prognosis); identifying possible high-penetrance subgroups (e.g. increased genotype-specific risk in pre-menopausal women); or conversely, identifying genetic subgroups with higher exposure-specific disease risk for prevention efforts (e.g increase environmental-specific risk for individuals with a particular genotype) (4-7). Furthermore, in the search for novel genes via GWAS, the modifying effects of environmental risk factors are not often taken into account; therefore, leveraging GxE may result in discovery of additional disease susceptibility loci (5, 8, 9). Despite interest in GxE, there are few agreed upon successes where the effect of exposure differs across genotypes (and vice versa). Numerous reasons have been suggested to contribute to the small number of successes including: the inherent low power of tests for GxE, complexity of measurement of environmental exposures and difficulty of incorporating temporality of environmental exposures, measurement error, limited range of genetic and/or environmental variation, scale dependence in the definition of statistical interaction, and lack of data on the biological consequences of most genetic variants (10-13).

The past few years have seen an emergence of new approaches, study designs, and statistical and analytical methods for exploring gene-environment interactions (GxE) in large-
scale human population studies. Further, new opportunities in this field, with respect to the incorporation of -omics and next-generation sequencing data and improvements in measures of environmental exposures implicated in complex disease outcomes, continue to be developed. Therefore, on October 17-18, 2014, National Institute of Environmental Health Sciences and National Cancer Institute held a workshop at the 64th Annual Meeting of the American Society of Human Genetics to explore these new approaches and tools for GxE interaction discovery. Based on the discussions, we prepared four papers that provide an update on: 1) the state of the science in analytical methods (14); 2) opportunities for incorporation of biological knowledge into GxE analyses (15); 3) advances in environmental exposure assessment in human populations (Chirag J. Patel, Department of Biomedical Informatics, Harvard Medical School, unpublished manuscript); and 4) lessons learned from GxE successes (Beate R. Ritz, Department of Epidemiology, Fielding School of Public Health, University of California Los Angeles, unpublished manuscript). In addition, this current paper develops some overarching themes and sets the stage for this series. As environmental factors may be modifiable, defining subpopulations of individuals most susceptible to environmental factors through GxE analysis may provide targets to improve public health. This idea is consistent with the goal for President Obama’s recently launched Precision Medicine initiatives at the National Institutes of Health (NIH) --to better understand how individual variability contributes to differences in response to treatment or prevention (16, 17).

**Analytical Methods**

Studies of GxE interaction require much larger sample sizes than studies targeting either genetic or environmental main effects alone (18). Further, when performing GxE on a genome-
wide scale, sometimes referred to as genome-wide interaction studies, sample size requirements are substantially further inflated to account for the multiple comparisons performed (5, 19). Therefore, a goal of GxE methods development has been to improve power to detect associations. As detailed in the accompanying manuscript (14), many different methods have been explored in the context of a case-control studies as alternatives to traditional GxE tests, including case-only (20), empirical Bayes (21), Bayes Model Averaging (22), joint tests (9, 23, 24), case parent approaches (25-27), and 2-step approaches (19, 23, 28-33). Other approaches include set-based methods, which combine multiple variants or GxEs and which may be particularly appropriate for studies of rare variants (34-38). In addition, several methods have been developed to analyze GxE for quantitative outcomes (39-46).

The large number of available methods, as well as novel software tools to support the application of these methods (29, 47, 48), create opportunities to better study GxE interactions in genome-wide settings. Researchers may therefore wonder which method to use for their studies. Several previous simulation studies suggest that none of these GxE methods is universally the most powerful approach (29, 30, 49-52). Therefore, decisions about the most appropriate approach depend on several considerations including the hypotheses to be tested, likely genetic architecture, study design attributes, and characteristics of the population being studied. Investigators should be cautious about applying multiple methods to their data without an a priori basis for choosing among the results, as simply picking those with the most "significant" findings to report would clearly be a biased strategy that could contribute to spurious associations and to what has been referred to as a “vibration of effects” (53, 54). Some of the new methods, however, provide flexible frameworks for combining multiple tests with an appropriate permutation procedure to evaluate the significance of the overall results (29, 30).
The collection of methods allows investigators to address specific scientific questions and offers new opportunities for studies of GxE in large populations.

**Functional Validation and Discovery**

Despite the recent success of GWAS at identifying risk loci, by design variants identified are not usually the causal variants, defined as the functional genetic variant that influences risk of disease and explains the association. Currently, the underlying biological mechanism contributing to disease risk is only known for a small proportion of these loci. Therefore, more research to functionally characterize risk loci is now being performed, providing opportunities by which GxE analyses may shed new insights into disease development (55). An understanding of the biological consequences of particular genetic differences could lead to specific mechanistic hypotheses, identifying relevant exposures to test and specifying relevant statistical models. As described in the accompanying manuscript (15), these approaches include utilizing functional annotations for discovery and validation, studying molecular phenotypes (e.g. epigenetics or gene expression) to improve GxE discovery, and leveraging *in vitro* and *in vivo* models for these studies.

Several large public databases [such as Encyclopedia of DNA Elements (ENCODE), Epigenomics Roadmap, Genotype-Tissue Expression (GTEx), and the Cancer Genome Atlas (TCGA)] have facilitated the functional annotation and interpretation of many genomic regions, which can be used to prioritize candidate GxE markers (30). Many disease-associated GWAS SNPs appear to be located in non-coding or regulatory regions which are often affected by environmental exposures (56-58). The ENCODE and Roadmap Epigenomics programs have helped to define many of the regulatory regions, and new tools developed by these programs and
others now allow functional annotation information, such as the genomic location of histone modification states, methylation patterns, transcription factor binding sites and DNAse hypersensitivity sites or other higher order chromosomal structural information, to be overlaid with GWAS results and could be integrated into GxE analyses (59-63). Projects like GTEx have greatly increased the compendium of putative biological functions of genetic variants. However, neither GTEx nor large-scale epigenomics projects provide information on effects of genetic and genomic functions across a range of environmental conditions. To explore genetic effects in response to environment, in vitro studies have now perturbed cells and recorded responses to various drugs, infections, and other exposures. Through use of intermediate molecular phenotypes such as gene expression, these efforts have demonstrated success in illustrating how an exposure may impact gene function, suggesting potential candidate genes or variants for GxE studies (64-68).

In addition to data resources, the use of population-based mouse resources (such as the Collaborative Cross, Diversity Outbred, and Hybrid Mouse Diversity Panel) and other appropriate mouse models have also been leveraged to assist in the discovery or replication of GxE interactions. These population-based variant enriched mouse resources have been designed to mimic the genetic diversity of human populations and can be used to replicate or inform GxE hypotheses by utilizing carefully controlled exposures in the mouse studies. Several recent examples have exemplified the power of these resources to map genetic variants related to susceptibility to environmental exposures (69, 70). Although both in vitro and model systems have led to potential mechanistic insights, linking of these to human populations remains challenging.
There are many approaches for incorporating biological knowledge to improve analytical methods (71) for GxE interaction in both the discovery and the validation phase. Incorporating functional annotation data and *a priori* biological information (such as metabolomics or gene expression data collected on individuals or knowledge on biological pathways) to inform data analytical GxE methods have aided in the discovery of new GxE findings in recent years (72). For example, Bayesian Variable Selection (73, 74), the Algorithm for Learning Pathways (75) and PEAK (72) are all methods that incorporate external biological information and properties of the dataset itself to increase power over agnostic approaches to detect interactions. Another approach is to use 2-stage modeling where functional annotations are used to prioritize variants (76, 77) for GxE studies. As one example, Biofilter was designed to build biologically plausible models of gene-gene and GxE interactions to test for associations based on biological features using biological knowledge from the public domain (76, 78). These types of filtering approaches are also being explored to prioritize environmental exposures by using databases such as the Comparative Toxicogenomics Database, which links exposures to genes (79). However, challenges still exist in linking environmental exposures into currently available ontological knowledge resources, though some investigators are beginning to navigate these challenges (80). Furthermore, all these databases and functional annotations depend on the quality and extent of existing biological knowledge (71).

**Environmental Exposures**

The complex realities of environmental exposures have long made measurement of exposures substantially more complicated than inherited genetic measurements (e.g. genotypes) and single nucleotide variants in particular; the technologies and approaches to incorporate
exposures into human population studies have therefore lagged behind genomics capabilities (11, 81). Assessing exposure impact must take into context not just the variety of exposures themselves (physical, often complex chemical mixtures, biological, and psychosocial) but also the source and place of exposure, the timing during a person’s life trajectory, the route of contact (skin, lung, diet), metabolism/excretion, and distribution in target tissues. All of these factors may impact the ultimate disease risk associated with environmental exposures. In addition, in the classic environmental exposure paradigm, studies may focus on measurements to capture internal versus external exposure, early markers of disease, or an ultimate biological response, which further adds to the complexity of exploring the impact of environmental exposures.

In recent years, however, exciting new opportunities have become available for environmental exposure assessment. The potential importance of examining the totality of internal and external exposures, referred to as the ‘exposome’, has been recognized (81, 82). Several recent commentaries described considerations for measurements of the exposome (83-86). Innovative technologies including activity monitors, improved sensors, global positioning systems, and Geographic Information Systems, which enable new and more detailed exposure measurements. Although issues of the timing of exposure measures persist and should be considered. Moreover, development of biological response markers for assessment of exposure, such as changes in gene expression, transcriptomic signatures, and DNA methylation profiles, has been useful for GxE discovery (87-90). Another opportunity is the exploration of environmental exposures in a more agnostic discovery-based fashion, similar to GWAS. These studies, termed environment wide association studies (EWAS) led to new discoveries of environmental factors associated with disease (91-94).
Key challenges and considerations remain associated with assessing environmental exposures in GxE studies, as detailed in the accompanying manuscript (Chirag J. Patel, Department of Biomedical Informatics, Harvard Medical School, unpublished manuscript) including how to: select most appropriate study designs, incorporate high throughput -omic measures (e.g. metagenome, metabolome) and sensor technologies into human population-based studies, assess long term exposure, integrate a variety of divergent external exposure and internal response data, and further advance statistical approaches to handle the dynamic nature of exposure data. We are now at the early stages of exploring what novel exposure assessment technologies can be appropriately applied to larger population studies most effectively. To this end, some two-stage study designs have been investigated (24, 95-98). Given the extreme cost of incorporating some sophisticated environmental measures in a large scale human population study, the question of what can be accomplished with dense (i.e. repeated measures of a marker or measurement of multiple analytes using an -omic platform) environmental measures on a subsample and extrapolating to a larger sample size (and whether simulations can demonstrate that this approach increases power to detect GxE) is currently being explored (49, 95, 99).

Several analytical methods have been developed for the unique considerations of exposure assessment. New statistical methods can adjust for exposure misclassification (which has been shown to lead to inflated type I errors and substantially reduced power) much better; these approaches should allow for obtaining greater power with smaller sample sizes. In addition, novel statistical methods have been developed to detect gene by longitudinal exposure interactions by taking into account long term time-varying exposures (100). Importantly, as researchers begin to combine exposure data to obtain larger sample sizes required for GxE research across studies, they have to address that exposures may have been measured using
different approaches or have very different distributions in and between populations such that exposure misclassification could produce spurious associations (14). There is also the challenge of exposure-related population stratification for studies relating to GxE interactions (101).

Meanwhile, multiple measures can sometimes increase power for detecting associations. For example, in a recent study, continuous monitoring was shown to reduce the sample size required in a clinical trial context (102).

**GxE Examples from Human Population Studies**

By examining GxE successes, it may be possible to improve the design of GxE studies for the future. Examples of GxE successes span from Mendelian-like traits (e.g. phenylketonuria) to complex diseases (NAT2 variants, smoking and bladder cancer) and response to therapies (HLA-B*1502 variant and carbamazepine induced Stevens Johnsons Syndrome) (Beate R. Ritz, Department of Epidemiology, Fielding School of Public Health, University of California Los Angeles, unpublished manuscript). In addition, several recent studies examined the use of polygenic risk scores, generated from common genetic variation, to assess the impact of environmental factors on individuals with low compared with higher genetic risk (103-106).

In the accompanying manuscript highlighting some of the most successful GxE interactions identified to date (Beate R. Ritz, Department of Epidemiology, Fielding School of Public Health, University of California Los Angeles, unpublished manuscript), several common themes have emerged including: the strength of focusing on metabolic pathways for a specific exposure, the utility of studying unique, highly or diversely exposed populations, the necessity of using high-quality exposure assessment methods, the need for large sample sizes, and the utility of model
systems to demonstrate genetic function when replication is challenging in population-based studies. These suggest important avenues for undertaking successful future research in GxE.

**Themes and Future Directions**

Inclusion of diverse populations may facilitate GxE research by improving power for discovery of casual genetic variants and environmental factors associated with disease. Trans-ethnic differences in the distribution of linkage disequilibrium can be leveraged to improve fine mapping to identify potential causal alleles (107-110). Combining admixture mapping with conventional GWAS may also facilitate discovery of novel loci (111). Using this later approach, novel loci were identified associated with total IgE levels (112) and asthma (113). Lastly, using geographically diverse populations might expand the distribution of the environmental exposure and thus increase power to detect interactions (13). Performing genetic studies on populations of diverse ancestry may improve our understanding of disease mechanisms and such studies are required to ensure all populations benefit equally from this research (114).

Replication is an essential component to genetic association studies, and the requirement for independent replication contributed to the success of GWAS (115, 116). However, replication and meta-analysis becomes challenging as GxE studies become sophisticated in analytical methods, exposure assessment, and incorporation of functional information. Differences in the underlying distribution of environmental exposures, genetic linkage disequilibrium (LD) structure, and genetic modifiers can reduce the power to detect the same level of interaction in independent studies. Moreover, an appropriate human replication study may not (yet) exist: in studies of a rare disease, genetic variant, or environmental exposure;
where exposures are unique to particular populations; or where the initial finding was obtained within a large consortium comprising all known studies of a specific outcome (12). As illustrated in the manuscripts describing GxE successes and incorporation of biological knowledge, in some situations functional studies could serve to provide support for initial GxE observations in absence of a suitable replication population. Moreover, as the field considers gene and pathway based approaches to study GxE, replication may become further complicated as different combinations of genes in different datasets may be observed in the interaction. Some have argued that replication requirements might be met if the underlying biological pathway is the same even if replication was not observed with the individual SNP or gene (15). More consideration of standards for replication, definitions of replication and alternative approaches for replication and verification of GxE results is needed.

Many exciting opportunities exist for studies of GxE. There is the emerging recognition that developmental exposures may lead to disease throughout life and efforts have focused on beginning to address how much of environmental exposure risk for many disease outcomes may be attributable to in utero exposures or other particularly vulnerable windows of susceptibility (childhood, adolescence, etc.). Successful integration of large volumes of diverse data types (including Geographic Information Systems, sensor, metabolomics and other omics data) will create the opportunity for generating unique insights. Epigenetics tools open up new opportunities to directly link environmental exposure to the genome and generate new exposure biomarkers (e.g. methylation of cancer specific genes associated with dietary folate and alcohol in colorectal cancer (117) or smoking exposure in lung cancer (118)) [for review (119, 120)]. Moreover, epigenomics, as well as other -omic technologies, may elucidate mechanisms by which exposures contribute to disease. The role of the microbiome as a key environmental risk
factor for many complex disease phenotypes is starting to be appreciated and extensively studied. In addition, molecular phenotype data creates opportunities to examine disease subtypes or more precisely classify disease. This may eventually reduce heterogeneity in studies and improve power to study GxE associations, assuming molecular characterizations are performed with the correct cell type, tissue, or appropriate surrogate tissue for the hypothesis being tested.

Additional areas of research may allow further advances in GxE discovery and replication. The field needs to determine how to best leverage experimental studies in animals or human cell lines to aid in discovering and functionally validating GxE interactions. Moreover, it is unclear how to best leverage existing family and twin based studies for examining GxE. In incorporating functional information into GxE studies, questions remain about the appropriate balance between using prior or external information versus the characteristics of the dataset being studied in building analytical models and appropriate methods for linking environmental exposures information into available biological knowledge databases that are usually focused on genes and pathways. In addition, since many GxE findings to date have modest effect sizes or have not been extensively replicated (11, 121), exploring the general question of when the effort of attempting to identify these complex types of interactions is worth it. Though even with modest effect sizes, if a GxE finding is sufficiently replicated in human populations and supported by other experimental data, this information could provide insights into possible disease mechanisms. Finally, given the reduced power to detect GxE combinations with present methods, approaches that examine higher order interactions should be taken on cautiously.

Despite many recent advances in analytical methods for GxE discovery and some validation in recent years, additional statistical methods are needed for studies of copy number and rare genetic variation, survival traits, analysis of trios, and meta-analysis and pooling in large
consortia. In addition, many of the assumptions about expected GxE findings are based on results from genetic simulation studies, but these expectations have not always directly correlated to GxE observations in real population studies. Therefore, the question remains whether simulation studies have been designed with realistic assumptions about the underlying genetic architecture of the traits and whether better simulation approaches are needed (122).

Extended collaboration and data sharing will also advance GxE research. Large epidemiological consortia, such as the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium, that have longitudinal measures of environmental exposures have been heavily leveraged in recent years as a way to examine repeated environmental exposures over time and attempt to incorporate cumulative and time-varying exposures into assessments of complex disease risk (123). There is also a need for further collaboration to allow validation of biomarkers in larger cohorts. Meta-analysis and pooling methodology and efforts will likely need to be advanced to have the power to detect GxE in rarer diseases. Standards are needed to describe the adequate criteria for identifying, reproducing, and reporting a GxE finding; a place to publish negative findings would allow researchers to avoid repeating failed experiments (11). There is also a need for greater integration and education with other fields to better design studies of GxE. Specifically, toxicology expertise will be needed to allow validation in experimental models of GxE discoveries. Lastly, compared to genomic data sharing, the sharing of environmental and epidemiological data has lagged behind. Some have suggested that an environmental data sharing policy mirroring the National Institutes of Health genomic data sharing policy could advance data sharing in the environmental health science fields. However, there are unique sensitivities and ethical issues related to the sharing of environmental data that must be considered, including participant confidentiality and privacy.
issues (i.e. environmental exposure data with global positioning systems information can allow specific identification of the sources of exposure) and legal/regulatory matters (i.e. regulatory reporting, remediation, and reform).

Researchers are exploring how to apply GxE findings to risk prediction studies as a possibility for targeted screening or intervention. Questions remain about the optimal approaches for risk prediction models, including how to integrate biomarkers and external exposures and how best to model the joint effects of genetic markers, biomarkers, and lifestyle and environmental exposures (124). Although most statistical methods for detecting GxE focus on identifying departures from a multiplicative relative risk model, the absence of multiplicative interactions will typically imply the presence of additive interaction (i.e., when there are marginal genetic and environmental effects). Additive interactions may have public health implications, as they suggest the difference in absolute risks between exposed and unexposed groups differs across genetically defined subgroups (103-106, 124). If an exposure causes disease, then an intervention to remove the exposure will prevent more cases in a genetically sensitive population than an in an equivalently sized genetically insensitive population.

Important challenges that remain include: determining whether the exposure in fact causes disease, developing effective interventions to change exposures, and evaluating whether targeted or population-level interventions optimize the risk: benefit trade-off. As with main effects, where it is well understood that observational findings of associations across individuals do not necessarily imply that an intervention to change exposure will change any individual’s outcomes, so an additive interaction does not necessarily imply that a genetically-targeted intervention would be a more effective prevention strategy. Modern methods of causal inference (125, 126) may be useful for estimating the causal difference in disease rates between genetically-targeted
and population-wide exposure interventions. Finally, the lessons and approaches for research into how the combination of genes and environment contribute to disease relates broadly to the studies of precision medicine and precision prevention. These types of studies may lead to insights for targeting prevention, intervention, or treatment in the future.

Acknowledgements

Author Affiliations:

Genes, Environment, and Health Branch, National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), Research Triangle Park, North Carolina (Kimberly McAllister); Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute (NCI), NIH, Bethesda, Maryland (Leah E. Mechanic); Department of Biomedical Data Science, Dartmouth College, Lebanon, New Hampshire (Christopher Amos); Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts; Centre de Bioinformatique, Biostatistique et Biologie Intégrative (C3BI), Institut Pasteur, Paris, France (Hugues Aschard); Center of Excellence in Environmental Toxicology and Penn SRP Center, Perelman School of Medicine, University of Pennsylvania Philadelphia, Pennsylvania; Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania Philadelphia, Pennsylvania (Ian A. Blair); Department of Biostatistics, Bloomberg School of Public Health, Department of Oncology, School of Medicine, Johns Hopkins University, Baltimore, Maryland (Nilanjan Chatterjee); Department of Preventive Medicine, University of Southern California, Los Angeles, California (David Conti); Department of Preventive Medicine, University of Southern California, Los Angeles, California (W. James Gauderman); Biostatistics and Biomathematics Program, Division of Public Health Sciences, Fred Hutchinson Cancer
Research Center, Seattle, Washington (Li Hsu); Division of Genome Sciences, National Human Genome Research Institute, NIH, Bethesda, Maryland (Carolyn M. Hutter); California Institute for Telecommunications and Information Technology, Qualcomm Institute, University of California San Diego, La Jolla California (Marta M. Jankowska); Department of Family Medicine and Public Health, University of California San Diego, La Jolla, California (Jacqueline Kerr); Department of Epidemiology, Harvard T.H. School of Public Health, Boston, Massachusetts (Peter Kraft); Departments of Genetics and Pathology, Stanford University School of Medicine, Stanford, California (Stephen B. Montgomery); Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, Michigan (Bhramar Mukherjee); Division of Cardiovascular Sciences, Prevention and Population Sciences Program, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland (George J. Papanicolaou); Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts (Chirag J. Patel); Department of Biochemistry and Molecular Biology, Center for Systems Genomics, The Pennsylvania State University, University Park, Pennsylvania; Biomedical and Translational Informatics, Geisinger Health System, Danville, Pennsylvania (Marylyn D. Ritchie); Department of Epidemiology, Fielding School of Public Health, University of California Los Angeles, Los Angeles, California (Beate R. Ritz); Department of Preventive Medicine, University of Southern California, Los Angeles, California (Duncan C. Thomas); Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas (Peng Wei); Department of Epidemiology and Biostatistics, University of California, San Francisco, California (John S. Witte)
We acknowledge NIEHS for providing funds to support this meeting. Research reported in this publication was supported by the National Cancer Institute, National Heart Lung and Blood Institute, National Human Genome Research Institute, and National Institute of Environmental Health Sciences of the National Institutes of Health, and the National Science Foundation under award numbers R21HG007687 to H.A.; R01CA140561, R01CA201407, and P01CA196569 to D.C.; R01CA189532, R01CA195789, and P01CA53996 to L.H.; R21CA169535 and R01CA179977 to J.K.; R21ES020811 and NSF DMS 1406712 to B.M.; R00ES023504 and R21ES025052 to C.J.P.; R01CA169122, R01HL116720 and R21HL126032 to P.W; and R01CA201358 to J.S.W. S.B.M. is supported by the National Institutes of Health through R01HG008150, R01MH101814, U01HG007436, and U01HG00908001. This work is funded, in part, under a grant with the Pennsylvania Department of Health (SAP 4100070267) to M.D.R. The authors thank the participants in the workshop “Current Challenges and New Opportunities for Gene-Environment Interaction Studies of Complex Diseases”. The Pennsylvania Department of Health specifically disclaims responsibility for any analyses, interpretations or conclusions. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The authors have no conflicts of interest to report.
References


