Title
Lessons Learned From Past Gene-Environment Interaction Successes.

Journal
American Journal of Epidemiology, 186(7)

ISSN
0002-9262

Authors
Ritz, Beate R
Chatterjee, Nilanjan
Garcia-Closas, Montserrat
et al.

Publication Date
2017-10-01

DOI
10.1093/aje/kwx230

Peer reviewed

Lessons Learned from Past Gene-Environment (GxE) Interaction Successes

Beate R. Ritz, Nilanjan Chatterjee, Montserrat Garcia-Closas, W. James Gauderman, Brandon L. Pierce, Peter Kraft, Caroline M. Tanner, Leah E. Mechanic, Kimberly McAllister

Running Title: Gene-Environment (GxE) Interaction Successes

Corresponding Author:

Beate R. Ritz, MD, PhD
Department of Epidemiology
Fielding School of Public Health
University of California Los Angeles
Los Angeles, CA, 90095
Phone: 310-825-8579
Fax: 310-206-6039
Email: britz@ucla.edu

Abbreviations

ALDH, aldehyde dehydrogenase; AS3MT, arsenite methyltransferase; CHRNA5, nicotinic receptor α5; CYP, cytochrome P-450; FeNO, nitric oxid; FTO, fat mass- and obesity; GxE, gene-environment; GWAS, genome-wide association study; G6PD, Glucose-6-phosphate dehydrogenase; GSTs, glutathione S-transferases; HLA-B, human leukocyte antigen-B; NAT2, N-acetyltransferase; NOS2, nitric oxide synthase 2; PD, Parkinson's Disease; PON1, paraoxonase 1; VKORC1, Vitamin K epoxide reductase complex subunit 1

Abstract

Genetic and environmental factors are both known to contribute to susceptibility to complex diseases. Therefore, the study of gene-environment (GxE) interactions has been a focus of research for several years. In this manuscript, select examples of GxE from the literature are described to highlight different approaches and underlying principles related to the success of these studies. These examples were broadly categorized as studies of: single metabolism genes, genes in complex metabolism pathways, ranges of exposure levels, functional approaches and model systems, and pharmacogenomics. Some studies illustrated the success of studying exposure metabolism for which candidate genes can be identified. Moreover, some GxE successes depended on the availability of high quality exposure assessment and longitudinal measures, study populations with a wide range of exposure levels, and the inclusion of ethnically and geographically diverse populations. In several examples, large population sizes were required to detect GxE interactions. Other examples illustrated the impact of accurately defining scale of the interactions (i.e. additive or multiplicative). Lastly, model systems and functional approaches provided insights into GxE in several examples. Future studies may benefit from these lessons learned.

Introduction

Genetic and environmental factors are both known to contribute to susceptibility to complex diseases. Studies of genetic variation range from hypothesis-driven studies examining a small number of candidate genes to more agnostic (i.e. hypothesis-free) surveys of variation across the entire genome, or genome-wide association studies (GWAS). GWAS leverage patterns of linkage disequilibrium with a high density of genetic markers to capture a large proportion of the common genetic variation in a population. Therefore, gene-environment (GxE) interactions, defined broadly as the interplay between gene(s) and environmental factor(s) as they affect some trait [discussed in (1)], are a focus of studies addressing chronic diseases such as neurodegeneration, cancer, and asthma, and more recently also pharmacogenomics applications; the former to better understand biological pathways to disease or identify subpopulations susceptible to specific exposures in human studies; the latter to contribute to 'precision medicine' and treatment plans tailored according to the genetic makeup of patients (2). We selected GxE examples that we knew to have been successful in reaching these aims illustrating different approaches and study designs which may have contributed to success.

The purpose of this paper is to describe a spectrum of approaches and underlying principles that have been successful for identifying GxE interactions to inform future studies. These success stories may be broadly categorized as studies of: single metabolism genes, complex metabolism pathways, broad ranges of exposures, functional approaches and model systems, and pharmacogenomics.

Single Metabolism Genes

Many of the genetic variants identified in GxE investigations are in metabolism genes and impact enzymatic function such that they may increase susceptibility to an environmental exposure and adversely affect health in exposed variant carriers. Recent efforts have focused on studying the genetics (or in some cases, epigenetics) of metabolism to identify single nucleotide polymorphisms that, by altering exposure metabolism, ultimately alter susceptibility to disease outcomes. Some of the oldest and best characterized GxE interactions with well-established biologic relevance and human health consequences are classic Mendelian genetic diseases of single metabolism genes that depend on the presence of common environmental (dietary or pharmacological) agents to adversely affect health (**Table 1, Lesson 1**), such as Phenylketonuria and Glucose-6-phosphate dehydrogenase (G6PD) deficiency. Phenylketonuria is caused by a defect in the gene encoding the enzyme phenylalanine hydroxylase which is needed to break down the amino acid phenylalanine. In children with Phenylketonuria, the resulting amino acid accumulation is responsible for severe intellectual and developmental disabilities, while dietary restrictions that eliminate phenylalanine intake can minimize or prevent adverse outcomes entirely (3). Similarly, G6PD deficiency is the most common human enzyme defect associated with neonatal jaundice and acute hemolytic anemia triggered by consumption of fava beans or treatment with antibiotic and antimalarial drugs (4). These disease phenotypes are entirely preventable and amenable to environmental interventions; the outcome only occurs if both the genetic and environmental factors are present (**Table 1, Lesson 2**).

Genes in Complex Metabolism Pathways

In contrast to Phenylketonuria and G6PD, many other metabolizing enzymes are biologically versatile and often redundant in their actions (e.g. (5)) which complicates relating

any single genotype to a phenotype of interest. Nevertheless, early GxE studies focused on polymorphic variants in many metabolism genes that impact enzymatic or metabolic function of proteins that activate or detoxify exogenous and endogenous toxins. Examples include members of the large microsomal oxidative cytochrome P-450 (*CYP*) superfamily of proteins (6), N-acetyltransferase 2 (*NAT2*) (7), and glutathione S-transferases (GSTs) (8) that are implicated in cancer (9), Parkinson's disease (PD) (10), and Alzheimer's disease (11). For example, long before the first familial PD gene was identified, the "poor metabolizer" enzymatic phenotype of the *CYP2D6* gene was the first PD candidate gene (12, 13) because the enzyme is active in the brain region linked to PD, metabolizes relevant endogenous neural compounds (14, 15), and inactivates neurotoxins known to cause Parkinsonism in animal models and humans (16). Many population studies have shown an increased risk of PD for *CYP2D6* poor metabolizers compared with all other metabolizer types (17), and some PD studies that include pesticide exposures also observed GxE interactions for the poor metabolizer variants of *CYP2D6* (18-21) (**Table 1, Lesson 1**).

Additionally, common variants of the paraoxonase 1 (*PON1)* gene act on the toxic oxon metabolite of organophosphate pesticides and are well characterized with regards to influence on enzyme activity in human serum (22, 23). Thus, in populations with chronic organophosphate exposure, carriers of *PON1* gene slow metabolizer variants have been shown to be at increased risk for PD (elderly) and developmental deficits (children) (**Table 1, Lesson 1**) (24). The PD studies were conducted among central California residents using records from the California Pesticide Use Reporting System to generate long-term OP pesticide estimates with sophisticated geographic information system tools (25). The neurodevelopmental studies were able to rely on biomarkers of exposure collected during short but relevant periods in pregnancy (26, 27). In

these studies, considering the temporality of exposure was important due to the potential impact on early life (i.e. neurodevelopmental outcomes) and later onset disease outcomes (i.e. PD) that needed to be examined to best characterize the interactions (**Table 1, Lesson 3**). Moreover, the exposure assessment approaches used in these studies were sophisticated and robust (**Table 1, Lesson 4**).

Another GxE example related to variation in metabolism genes is one of the most well-established GxE interactions in cancer -- the association of genetic variation in N-acetyltransferase-2 (*NAT2*), smoking, and risk of bladder cancer (28) (**Table 1, Lesson 1**). *NAT2* catalyzes metabolism of aromatic monamines which are known bladder carcinogens found in cigarette smoke. Several common genetic variants in *NAT2* are related to reduced enzyme activity (29) and segregate populations into rapid, intermediate, and slow acetylation phenotypes which influence the ability to detoxify aromatic amines. Since tobacco smoking is a strong risk factor for bladder cancer and aromatic amines found in tobacco smoke are known bladder carcinogens, reduced detoxification capacity was hypothesized to increase susceptibility. Indeed, studies in different populations consistently demonstrated that slow acetylation activity increases risk of bladder cancer among smokers, but not among never smokers (7, 30-34). Many studies utilized candidate gene approaches focusing on the hypothesis related to the role of *NAT2* in metabolism (7, 32-34). A genome-wide interaction analysis of smoking and bladder cancer, however, observed interactions with different single nucleotide polymorphisms depending on whether the interaction was evaluated on an additive or multiplicative scale (30), highlighting the importance of defining the scale of measurement discussed in manuscript 2 of this series (1) and previously (28, 35, 36) (**Table 1, Lesson 5).** Interestingly, the multiplicative interaction between *NAT2* and smoking, even though it is supported by strong prior knowledge, did not reach a

genome-wide significance threshold. The authors estimated it would require over 15,000 cases with a 1:2 ratio of cases:controls to reach the statistical significance threshold, i.e. large sample sizes are required to discover GxE interactions without strong priors due in part to the stringent significance thresholds required for agnostic studies (30). Meanwhile, this interaction was observed in previous candidate gene studies with 1100-3000 cases illustrating the power of hypothesis-driven studies compared to GWAS approaches when prior knowledge exists (**Table 1, Lesson 5, 6).**

The well-established interaction between a variant in the aldehyde dehydrogenase 2 (*ALDH2*) gene and alcohol on risk of esophageal squamous-cell carcinoma (37-41) highlights several other considerations. Alcohol is oxidized to form acetaldehyde, a carcinogen. *ALDH2* detoxifies acetaldehyde to acetate. The *ALDH2*2* allele slows this detoxification process (**Table 1, Lesson 1**). The obvious hypothesis that an increase in risk due to alcohol consumption will be larger among individuals who carry the *ALDH2*2* allele has been borne out in observational studies (38-41). Originally studied as a candidate gene, the *ALDH2*2* allele has also been 'rediscovered' via GWAS, using both a marginal approach (testing for association without considering effect modification by alcohol intake) and using a GxE interaction approach (41). There are two particularities to the *ALDH2* example worth noting. First, the *ALDH2*2* allele is common in East Asian populations but quite rare in European-ancestry populations. This underscores the importance of conducting studies in ethnically diverse populations—not only can this increase the diversity of environmental exposures, it can increase the diversity of genetic exposures (**Table 1, Lesson 7**). Second, *ALDH2*2* is associated with exposure. Individuals who carry an *ALDH2*2* allele experience an unpleasant flushing reaction to alcohol and are less likely to drink regularly or heavily (37). This gene-environment correlation has implications for

downstream analysis. Tests that assume that genotype and environmental exposure are independent—such as the case-only test, a method to test association between genetic and environmental factors within exposed and unexposed cases only —can be more powerful than tests that do not make this assumption *when the assumption holds* (42). Although this assumption may be reasonable for many (or most) tested variants (42), when it is violated, tests assuming gene-environment independence can have inflated Type I error or decreased power. In the case of *ALDH2*2*, because the gene-environment correlation and GxE interaction act in opposite directions, the case-only test failed to detect the *ALDH2*2*-alcohol interaction at a nominal level of association in a study where the standard logistic regression test of interaction including controls was highly significant (43). As this example illustrates, the most appropriate method or approach for analyzing GxE interactions can be highly dependent on an understanding of underlying assumptions and correlations between risk factors. As discussed in more detail in a companion paper,, no single GxE method is universally the most powerful approach and efficiency will depend on the hypothesis tested and the underlying true GxE model (1, 44) (**Table I, Lesson 8**).

Variation of Exposure Levels

Low exposure variability may be one key factor for not being able to identify GxE interactions (45). A wide range of exposure levels among study participants provides greater statistical power for identifying GxE interactions (1) and an opportunity to contrast different models (e.g. linear, threshold, age-specific) to understand how exposure impacts the gene-trait relationship.

A compelling example of the benefits of targeting a population with both high and low exposure scenarios comes from a study of interaction between physical activity and the fat mass-

and obesity associated (*FTO*) gene on waist circumference conducted in a multi-center study in India (46). The study included two sites, one in northern India (New Delhi) and the other in southern India (Trivandrum), each including close to 500 individuals. The population in Delhi had low levels of physical activity, similarly to what is typically observed in Caucasian populations. In contrast, the population in Trivandrum had a wide range of physical activity levels and included individuals with high or low activity. The original association between *FTO* genetic variation and obesity, generated primarily in Caucasian populations (47-49), was replicated in the Delhi but not the Trivandrum population. However, in Trivandrum, an interaction was detected between physical activity and *FTO* – the association between *FTO* variant and obesity was strongest in individuals with low physical activity and diminished gradually with increasingly higher levels of physical activity. This pattern of interaction has been replicated in studies conducted in Caucasian populations (50). However, due to a narrower range of physical activity a much larger sample size (N>200,000) was required for this interaction to achieve statistical significance (**Table 1, Lesson 6**). This example demonstrates advantages of having robust measures of environmental factors with a wide range of exposure levels for identifying GxE interactions, with strong variation in exposure increasing the power to detect GxE. Studying geographically, culturally, or sociologically diverse populations may make it more likely to observe variations in exposure (**Table 1, Lessons 4, 7, 9**). Another study showed that the influence of a common *FTO* variant on body mass index varies across birth cohorts, calendar time periods, and life cycles (51) illustrating the importance that temporal considerations may have for GxE interaction studies and demonstrating that global or local environmental changes over time can modify the observed allelic penetrance of genetic risk factors for complex traits (**Table 1, Lesson 3**).

Another approach to achieving an adequate range of exposures is to examine a highly exposed population. Long term exposure assessment on a study population highly exposed to a specific agent from the environment may provide unique insights for characterizing GxE if the population is well characterized longitudinally. Inorganic arsenic is a known human carcinogen (52) and the natural or man-made contamination of ground water used as drinking water in several regions across the globe makes this exposure a serious global health issue (53, 54), particularly in Bangladesh where >57 million people are exposed at levels exceeding the WHO recommended limit (55, 56). A GWAS-approach was used to identify genetic polymorphisms associated with an "arsenic metabolism efficiency" phenotype, and these variants were found to affect risk of arsenical skin lesions (57). Arsenic metabolism efficiency can be measured in urine as ratios of arsenic metabolites to total arsenic. Dimethylarsinic acid, the end metabolite, is most readily excreted in urine, and individuals with high dimethylarsinic acid % are viewed as more efficient metabolizers with lower risk for arsenic toxicity (58-62). The recent GWAS (57) identified two independent association signals for dimethylarsinic acid % in a 10q24.32 region containing *AS3MT*, a gene involved in arsenic methylation, consistent with several candidate gene studies (reviewed in (63)). The low-efficiency alleles at these two single nucleotide polymorphisms were independently associated with increased risk for arsenical skin lesions. Furthermore, the association between arsenic exposure and skin lesion risk was weaker among individuals with high-efficiency 10q24.32 genotypes than those with low-efficiency genotypes (64). This example illustrates the strength of using a highly exposed population with sufficient variation in exposure (**Table 1, Lesson 7**) and using high-quality exposure measures to identify GxE (**Table 1, Lesson 9**). This example further illustrates the strength of studying metabolism pathways of an exposure to identify genetic variants that impact disease (**Table 1, Lesson 1**).

Model Systems and Functional Approaches

Replication has been a cornerstone of genetic association studies, and the requirement for independent replication contributed to the success of GWAS (65, 66). In examples described above, particularly *NAT2* x smoking with bladder cancer, *FTO* and physical activity with body mass index, and *ALDH2* x alcohol with esophageal cancer, interactions were replicated multiple times in different populations. However, as discussed in the accompanying manuscripts and previously, there are many challenges to replication in the GxE context and what constitutes appropriate replication in the GxE context is currently being debated (28, 44, 67, 68) and (Chirag J. Patel, Department of Biomedical Informatics, Harvard Medical School, unpublished manuscript). Functional approaches can complement and support population based epidemiology studies by providing potential mechanistic insights to observed findings (**Table 1, Lesson 10**) (69) and are described in a companion manuscript (68). These approaches include model systems, *in vitro* experiments, and biomarker measurements. They may also provide additional evidence for an interaction in situations where replication in another independent population is not possible due to lack of availability of an appropriate replication population, such as in studies of a rare disease or environmental exposure (28, 67).

An example where functional approaches helped to validate genetic associations that were difficult to resolve with human data involves chronic lead exposure and genetic polymorphisms affecting lead processing and excretion functions (70-72). Although genetically driven variations in human susceptibility to adverse health effects from lead toxicity is a well-appreciated phenomenon (73, 74), studies trying to establish the genetics of human susceptibility

have been challenging due to the variety of clinical symptoms elicited by lead toxicity. Moreover, studies with chronic low levels of lead exposures require long-term exposure assessments that account for life cycle susceptibility such as during pregnancy and early childhood, which are difficult in human populations. Using *Drosophila melanogaster*, researchers used mutants to assess functional causality of candidate genes and were able to identify a genetic network related to lead susceptibility, building upon known genes previously identified in human GWAS (75) (**Table 1, Lesson 10**).

Insights into the underlying mechanisms related to the *FTO* gene and interaction with physical activity for obesity phenotypes was provided in a series of mechanistic studies. Energy balance is known to be modulated by both food consumption and physical activity as well as by the dissipation of energy as heat through constitutive heat generation (thermogenesis) in mitochondria-rich brown adipocytes in brown fat and through inducible thermogenesis in beige adipocytes in white fat. Thermogenesis is triggered in part by response to exercise and partially controlled by mitochondria. Functional studies have recently shown that one of the common *FTO* alleles associated with obesity phenotypes can repress mitochondrial thermogenesis in adipocyte precursor cells (76, 77). This leads to a developmental shift from energy-dissipating beige adipocytes to energy-storing white adipocytes with repression of basal mitochondrial respiration and increased lipid storage. These functional studies provide some biological evidence for the interaction of *FTO* and physical activity in generating obesity. These predictions were further functionally validated with knockdown and overexpression of the *FTO* gene and other regulators in human patient tissue samples and mice models (76, 77) (**Table 1, Lesson 10**).

Furthermore, functional studies of Parkinson's disease (PD) helped elucidate findings from human genetic association studies. Pesticide exposure was suggested as an environmental risk factor for PD, though the mechanism was unknown. Aldehyde dehydrogenase (ALDH) plays a key role in neuronal protection by metabolizing biogenic amine-related aldehydes, e.g., 3,4-dihydroxyphenylacetaldehyde, and by protecting neurons against aldehyde- and oxidative stress-related neurotoxicity (78) (**Table 1, Lesson 1**). Therefore, researchers used an ex vivo model system to identify several pesticides which inhibited the enzyme activity of ALDH (**Table 1, Lesson 10**). These same pesticides were associated with an increased risk of PD in a population-based study, and genetic variation in the *ALDH2* appeared to modulate PD risk due to these pesticide exposures (79).

Identifying a plausible biological mechanism using biomarkers can also help validate human population study findings, as illustrated by GxE interactions in asthma. Exhaled nitric oxide (FeNO) levels are known biomarkers of airway inflammation that are predictive of childhood asthma (80). Researchers found that common inducible nitric oxide synthase 2 (*NOS2*) promoter haplotypes combined with residential traffic-related exposure appeared to interact to affect exhaled nitric oxide (FeNO) levels in children, presumably because *NOS2* is induced by environmental exposures (81). Previously, common genetic variants and promoter haplotypes of *NOS2* were associated with childhood exhaled FeNO values (80). Moreover, exposure to residential traffic and allergens were independently associated with elevated FeNO levels (81). The discovery of this GxE interaction (*NOS2* promoter haplotypes x traffic exposure) benefited from a large, well-characterized population with substantial variation in exposure levels (**Table 1, Lessons 4 and 7**). In addition, higher FeNO levels were associated with elevated *NOS2* mRNA in the bronchial epithelium of asthmatics after allergen exposure

(82). Although this GxE finding needs to be replicated, the biomarker study correlating higher FeNO levels with higher NOS2 expression suggests a plausible biological mechanism of how ubiquitous air pollutants and genetic variation might impact a biological pathway relevant to inflammation that could contribute to asthma (**Table 1, Lesson 10**).

Pharmacogenomics

There are several lessons to be learned from pharmacogenomic GxE studies that may be applied more broadly to GxE studies of complex diseases. Pharmacogenomics, which specifically examines the role of genetic variation in various drug response phenotypes (83-85), can utilize either targeted or untargeted GxE approaches with the drug as the environmental exposure and drug response phenotype as the outcome (86). Variations in drug response may cause some individuals to require a higher dose while others require a lower dose because of increased sensitivity or adverse side effects. For example, common genetic variants in Cytochrome P450 2C9 (*CYP2C9*) and Vitamin K epoxide reductase complex subunit 1 (VKORC1) genes contribute to variability in patients' responses to treatment with the anticoagulant warfarin, explaining as much as 18-30% of the response variability observed in European populations (87) (**Table 1, Lesson 2**). Notably, many successes in pharmacogenomics have been observed, despite the challenges of small population sizes and rarity of adverse events (84, 85, 88), due to the fact that the environmental agent (i.e. the drug), is known, easy to measure, and is often associated with a well-defined outcome phenotype, such as lowering blood pressure (84). In addition, for many pharmaceuticals, the mechanism of action and metabolic pathways are well-understood, making targeted studies with a prior hypothesis more successful than agnostic

GWAS studies, particularly given the usual small population sizes of these studies, and findings from such studies easily interpretable (**Table 1, Lesson 1**).

An important lesson from pharmacogenomics studies relates to the importance of studying diverse populations since considerable variation in drug response across ethnicities have been observed. In many cases, this variation is due to frequencies of genetic variants depending on population ancestry (84). For example, in East and South-East Asian populations, a strong association of carbamazepine-induced Stevens Johnsons Syndrome was reported for the human leukocyte antigen-B (HLA-B)*1502 allele (OR 84.75; 95% CI 42.53-168.91; P=8.96 x 10^{-5}), while no associations were observed in Japanese or Caucasian patients (89). To date, a majority of GWAS and pharmacogenomics studies have been conducted in populations of European descent. Performing genetic studies on populations of diverse ancestry will likely provide further insights into disease mechanisms and ensure that all populations derive benefits from pharmacogenomics research (90) (**Table 1, Lesson 7**).

GxE findings in pharmacogenomics may also be applied to disease prevention, such as smoking cessation. Evidence suggests that both nicotinic receptor α5 subunit (*CHRNA5*) and cytochrome P450 2A6 (*CYP2A6*) genotypes influence smoking cessation success and response to pharmacotherapy. In a large smoking cessation trial, the effectiveness of smoking cessation pharmacotherapy and medication efficacy was dependent on *CHRNA5* haplotype (91). Similar pharmacogenomic interactions were observed in patient responses to nicotine replacement therapy with *CHRNA5* (92) and *CYP2A6* genetic variants (93). An additional study reported that the prescription medication varenicline was more efficacious than nicotine patches and depended on the *CYP2A6* genotypes for slow metabolizers, while the effect of the common drug bupropion

on smoking relapse did not seem to be affected by *CYP2A6* genotype-driven nicotine metabolism (93, 94). These studies support the notion that personalized smoking cessation intervention based on genotype may increase the effectiveness of such treatments (**Table 1, Lesson 2**).

In another example of GxE interactions in pharmacogenomics related to disease prevention, researchers performed an agnostic genome-wide GxE gene discovery study in colorectal cancer patients with regular use of aspirin or non-steroidal anti-inflammatory drugs or both medications. Use of aspirin and/or non-steroidal anti-inflammatory drugs was associated with reduced risk of colorectal cancer in individuals with *MGST1* TT genotypes and higher risk among those with the TA or AA genotypes (89). Meanwhile, regular use of aspirin and/or non-steroidal anti-inflammatory drugs was associated with lower risk of colorectal cancer among individuals with *IL16* AA genotypes but not with the less common genotypes. These results may have implications for targeting populations at risk of colorectal cancer for specific intervention efforts such as treatment with non-steroidal anti-inflammatory drugs and/or aspirin based on genetic information (**Table 1, Lesson 2**). To detect these interactions, the investigators combined data from 10 observational studies for a total of 8634 cases and 8533 controls. Even with this large sample size, only a few interactions were observed (**Table 1, Lesson 6**).

Conclusions

The characteristics of the above examples illustrate several important lessons for GxE research. First, studying variants that are known to disrupt exposure metabolism is a promising strategy for identifying disease-related variants that interact with exposure. Other pathways where mechanisms of exposure action are well-understood (e.g. pharmacogenomics) may also be successful approaches. Second, studying GxE in human studies designed to characterize a

specific exposure (such as arsenic or specific pesticides) over an extended period and in a large population will provide opportunities to utilize high-quality exposure measures, to study a wide range of exposure levels, and to examine longitudinal measures of exposure. Importantly, using carefully collected and comprehensive exposure data with a wide variation among study participants will increase statistical power for GxE detection. Even with high quality exposure assessment, many of these studies required large population sizes for the GxE discovery. In addition, GxE research should include diverse populations representing many geographic areas, cultures, and ethnicities. Finally, functional studies including model systems, laboratory studies or biomarkers measured in human tissues may lead to valuable insights to GxE findings, complimenting large population based epidemiology findings.

Acknowledgements

Author Affiliations:

Department of Epidemiology, Fielding School of Public Health, University of California Los Angeles, Los Angeles, California (Beate R. Ritz); Department of Biostatistics, Bloomberg School of Public Health, Department of Oncology, School of Medicine, Johns Hopkins University, Baltimore, Maryland (Nilanjan Chatterjee); Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institute of Health, Bethesda, Maryland (Montserrat Garcia-Closas); Department of Preventive Medicine, University of Southern California, Los Angeles, California (W. James Gauderman); Department of Public Health Sciences, Department of Human Genetics, Comprehensive Cancer Center, University of Chicago, Chicago, Illinois (Brandon L. Pierce); Department of Epidemiology, Harvard T.H. School of Public Health, Boston, Massachusetts (Peter Kraft); San Francisco Veterans Affairs Medical Center and Department of Neurology, University of California--San Francisco, San Francisco, California (Caroline M. Tanner); Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, NIH, Bethesda, Maryland (Leah E. Mechanic); Genes, Environment, and Health Branch, National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), Research Triangle Park, North Carolina (Kimberly McAllister)

Research reported in this publication was supported by National Institute of Environmental Health Sciences and the National Cancer Institute of the National Institutes of Health under award numbers P01CA196569, R01CA201407, P30ES07048, and R21ES024844 to W.J.G.; and R01ES023834 and R21ES024834 to B.L.P. 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

1. Gauderman WJ, Mukheerjee B, Aschard H, et al. Update on the State of the Science for Analytical Methods. *Am J Epidemiol* 2017;in press.
2. Thomas D. Gene-environment-wide association studies: emerging approaches. *Nature reviews Genetics* 2010;11(4):259-272.
3. National Institutes of Health Consensus Development Panel. National Institutes of Health Consensus Development Conference Statement: Phenylketonuria: Screening and Management, October 16–18, 2000. *Pediatrics* 2001;108(4):972-982.
4. Cappellini MD, Fiorelli G. Glucose-6-phosphate dehydrogenase deficiency. *Lancet* 2008;371(9606):64-74.
5. Bernhardt R. Cytochromes P450 as versatile biocatalysts. *J Biotechnol* 2006;124(1):128-145.
6. Blackburn HL, Ellsworth DL, Shriver CD, et al. Role of cytochrome P450 genes in breast cancer etiology and treatment: effects on estrogen biosynthesis, metabolism, and response to endocrine therapy. *Cancer Causes Control* 2015;26(3):319-332.
7. García-Closas M, Malats N, Silverman D, et al. NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *The Lancet* 2005;366(9486):649-659.
8. Hayes JD, Pulford DJ. The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit Rev Biochem Mol Biol* 1995;30(6):445-600.
9. Zhou LP, Luan H, Dong XH, et al. Genetic variants of CYP2D6 gene and cancer risk: a HuGE systematic review and meta-analysis. *Asian Pac J Cancer Prev* 2012;13(7):3165-3172.
10. Alonso-Navarro H, Jimenez-Jimenez FJ, Garcia-Martin E, et al. Genomic and pharmacogenomic biomarkers of Parkinson's disease. *Curr Drug Metab* 2014;15(2):129-181.
11. Allen M, Zou F, Chai HS, et al. Glutathione S-transferase omega genes in Alzheimer and Parkinson disease risk, age-at-diagnosis and brain gene expression: an association study with mechanistic implications. *Mol Neurodegener* 2012;7:13.
12. Barbeau A, Cloutier T, Roy M, et al. Ecogenetics of Parkinson's disease: 4-hydroxylation of debrisoquine. *Lancet* 1985;2(8466):1213-1216.
13. Riedl AG, Watts PM, Jenner P, et al. P450 enzymes and Parkinson's disease: the story so far. *Mov Disord* 1998;13(2):212-220.
14. Bromek E, Haduch A, Daniel WA. The ability of cytochrome P450 2D isoforms to synthesize dopamine in the brain: An in vitro study. *Eur J Pharmacol* 2010;626(2-3):171-178.
15. Hiroi T, Imaoka S, Funae Y. Dopamine formation from tyramine by CYP2D6. *Biochem Biophys Res Commun* 1998;249(3):838-843.
16. Langston JW, Ballard P, Tetrud JW, et al. Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis. *Science* 1983;219(4587):979-980.
17. Lu Y, Peng Q, Zeng Z, et al. CYP2D6 phenotypes and Parkinson's disease risk: a meta-analysis. *J Neurol Sci* 2014;336(1-2):161-168.
18. Dick FD, De Palma G, Ahmadi A, et al. Gene-environment interactions in parkinsonism and Parkinson's disease: the Geoparkinson study. *Occup Environ Med* 2007;64(10):673-680.
19. Elbaz A, Levecque C, Clavel J, et al. CYP2D6 polymorphism, pesticide exposure, and Parkinson's disease. *Ann Neurol* 2004;55(3):430-434.
20. Mellick GD. CYP450, genetics and Parkinson's disease: gene x environment interactions hold the key. *J Neural Transm Suppl* 2006(70):159-165.

21. Singh NK, Banerjee BD, Bala K, et al. Gene-gene and gene-environment interaction on the risk of Parkinson's disease. *Curr Aging Sci* 2014;7(2):101-109.
22. Costa LG, Cole TB, Vitalone A, et al. Measurement of paraoxonase (PON1) status as a potential biomarker of susceptibility to organophosphate toxicity. *Clin Chim Acta* 2005;352(1-2):37-47.
23. O'Leary KA, Edwards RJ, Town MM, et al. Genetic and other sources of variation in the activity of serum paraoxonase/diazoxonase in humans: consequences for risk from exposure to diazinon. *Pharmacogenet Genomics* 2005;15(1):51-60.
24. Gonzalez-Alzaga B, Lacasana M, Aguilar-Garduno C, et al. A systematic review of neurodevelopmental effects of prenatal and postnatal organophosphate pesticide exposure. *Toxicol Lett* 2014;230(2):104-121.
25. Goldberg DW, Wilson JP, Knoblock CA, et al. An effective and efficient approach for manually improving geocoded data. *Int J Health Geogr* 2008;7:60.
26. Engel SM, Berkowitz GS, Barr DB, et al. Prenatal organophosphate metabolite and organochlorine levels and performance on the Brazelton Neonatal Behavioral Assessment Scale in a multiethnic pregnancy cohort. *Am J Epidemiol* 2007;165(12):1397-1404.
27. Eskenazi B, Kogut K, Huen K, et al. Organophosphate pesticide exposure, PON1, and neurodevelopment in school-age children from the CHAMACOS study. *Environ Res* 2014;134:149-157.
28. Hutter CM, Mechanic LE, Chatterjee N, et al. Gene-environment interactions in cancer epidemiology: a National Cancer Institute Think Tank report. *Genet Epidemiol* 2013;37(7):643-657.
29. Hein DW. Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis. *Mutat Res* 2002;506-507:65-77.
30. Figueroa JD, Han SS, Garcia-Closas M, et al. Genome-wide interaction study of smoking and bladder cancer risk. *Carcinogenesis* 2014;35(8):1737-1744.
31. Garcia-Closas M, Rothman N, Figueroa JD, et al. Common genetic polymorphisms modify the effect of smoking on absolute risk of bladder cancer. *Cancer Res* 2013;73(7):2211-2220.
32. Gu J, Liang D, Wang Y, et al. Effects of N-acetyl transferase 1 and 2 polymorphisms on bladder cancer risk in Caucasians. *Mutat Res* 2005;581(1-2):97-104.
33. Moore LE, Baris DR, Figueroa JD, et al. GSTM1 null and NAT2 slow acetylation genotypes, smoking intensity and bladder cancer risk: results from the New England bladder cancer study and NAT2 meta-analysis. *Carcinogenesis* 2011;32(2):182-189.
34. Yuan JM, Chan KK, Coetzee GA, et al. Genetic determinants in the metabolism of bladder carcinogens in relation to risk of bladder cancer. *Carcinogenesis* 2008;29(7):1386-1393.
35. Thomas D. *Statistical methods in genetic epidemiology*. New York, NY: Oxford University Press; 2004.
36. Garcia-Closas M, Jacobs K, Kraft P, et al. Analysis of epidemiologic studies of genetic effects and gene-environment interactions. *IARC Sci Publ* 2011(163):281-301.
37. Brooks PJ, Enoch MA, Goldman D, et al. The alcohol flushing response: an unrecognized risk factor for esophageal cancer from alcohol consumption. *PLoS Med* 2009;6(3):e50.
38. Cui R, Kamatani Y, Takahashi A, et al. Functional variants in ADH1B and ALDH2 coupled with alcohol and smoking synergistically enhance esophageal cancer risk. *Gastroenterology* 2009;137(5):1768-1775.
39. Hiyama T, Yoshihara M, Tanaka S, et al. Genetic polymorphisms and esophageal cancer risk. *Int J Cancer* 2007;121(8):1643-1658.
40. Wu C, Hu Z, He Z, et al. Genome-wide association study identifies three new susceptibility loci for esophageal squamous-cell carcinoma in Chinese populations. *Nat Genet* 2011;43(7):679-684.

41. Wu C, Kraft P, Zhai K, et al. Genome-wide association analyses of esophageal squamous cell carcinoma in Chinese identify multiple susceptibility loci and gene-environment interactions. *Nat Genet* 2012;44(10):1090-1097.
42. Mukherjee B, Ahn J, Gruber SB, et al. Testing gene-environment interaction in large-scale case-control association studies: possible choices and comparisons. *Am J Epidemiol* 2012;175(3):177-190.
43. Wu C, Chang J, Ma B, et al. The case-only test for gene-environment interaction is not uniformly powerful: an empirical example. *Genet Epidemiol* 2013;37(4):402-407.
44. McAllister K, Mechanic LE, Amos CI, et al. Current Challenges and New Opportunities for Gene-Environment Interaction Studies of Complex Diseases. *Am J Epidemiol* 2017;in review.
45. Kraft P, Aschard H. Finding the missing gene-environment interactions. *Eur J Epidemiol* 2015;30(5):353-355.
46. Moore SC, Gunter MJ, Daniel CR, et al. Common genetic variants and central adiposity among Asian-Indians. *Obesity (Silver Spring)* 2012;20(9):1902-1908.
47. Dina C, Meyre D, Gallina S, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet* 2007;39(6):724-726.
48. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007;316(5826):889-894.
49. Scuteri A, Sanna S, Chen WM, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS genetics* 2007;3(7):e115.
50. Kilpelainen TO, Qi L, Brage S, et al. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med* 2011;8(11):e1001116.
51. Rosenquist JN, Lehrer SF, O'Malley AJ, et al. Cohort of birth modifies the association between FTO genotype and BMI. *Proc Natl Acad Sci U S A* 2015;112(2):354-359.
52. International Agency for Research on Cancer. *Some Drinking-water Disinfectants and Contaminants, including Arsenic: Arsenic in Drinking-water*. Lyon, France: IARC, 2004. (IARC monographs on the evaluation of carcinogenic risks to humans vol 84)
53. Brinkel J, Khan MH, Kraemer A. A systematic review of arsenic exposure and its social and mental health effects with special reference to Bangladesh. *Int J Environ Res Public Health* 2009;6(5):1609-1619.
54. Naujokas MF, Anderson B, Ahsan H, et al. The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environ Health Perspect* 2013;121(3):295-302.
55. British Geological Survey and the Department of Public Health Engineering. *Arsenic contamination of groundwater in Bangladesh. British Geological Survey Technical Report WC/00/19*. Keyworth, England, 2001. (British Geological Survey)
56. World Health Organization. Arsenic. 2016. (http://www.who.int/mediacentre/factsheets/fs372/en/). (Accessed 3/29/2017).
57. Pierce BL, Kibriya MG, Tong L, et al. Genome-wide association study identifies chromosome 10q24.32 variants associated with arsenic metabolism and toxicity phenotypes in Bangladesh. *PLoS genetics* 2012;8(2):e1002522.
58. Kile ML, Hoffman E, Rodrigues EG, et al. A pathway-based analysis of urinary arsenic metabolites and skin lesions. *Am J Epidemiol* 2011;173(7):778-786.

59. Ahsan H, Chen Y, Kibriya MG, et al. Arsenic metabolism, genetic susceptibility, and risk of premalignant skin lesions in Bangladesh. *Cancer Epidemiol Biomarkers Prev* 2007;16(6):1270-1278.
60. Lindberg AL, Rahman M, Persson LA, et al. The risk of arsenic induced skin lesions in Bangladeshi men and women is affected by arsenic metabolism and the age at first exposure. *Toxicol Appl Pharmacol* 2008;230(1):9-16.
61. Valenzuela OL, Borja-Aburto VH, Garcia-Vargas GG, et al. Urinary trivalent methylated arsenic species in a population chronically exposed to inorganic arsenic. *Environ Health Perspect* 2005;113(3):250-254.
62. Gao J, Yu J, Yang L. Urinary arsenic metabolites of subjects exposed to elevated arsenic present in coal in Shaanxi Province, China. *Int J Environ Res Public Health* 2011;8(6):1991-2008.
63. Agusa T, Fujihara J, Takeshita H, et al. Individual variations in inorganic arsenic metabolism associated with AS3MT genetic polymorphisms. *Int J Mol Sci* 2011;12(4):2351-2382.
64. Pierce BL, Tong L, Argos M, et al. Arsenic metabolism efficiency has a causal role in arsenic toxicity: Mendelian randomization and gene-environment interaction. *Int J Epidemiol* 2013;42(6):1862-1871.
65. Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype associations. *Nature* 2007;447(7145):655-660.
66. Kraft P, Zeggini E, Ioannidis JP. Replication in genome-wide association studies. *Stat Sci* 2009;24(4):561-573.
67. Mechanic LE, Chen HS, Amos CI, et al. Next generation analytic tools for large scale genetic epidemiology studies of complex diseases. *Genet Epidemiol* 2012;36(1):22-35.
68. Ritchie MD, Davis JR, Aschard H, et al. Incorporation of Biological Knowledge into the Study of GxE. *Am J Epidemiol* 2017;in press.
69. Boffetta P, Winn DM, Ioannidis JP, et al. Recommendations and proposed guidelines for assessing the cumulative evidence on joint effects of genes and environments on cancer occurrence in humans. *Int J Epidemiol* 2012;41(3):686-704.
70. Jhun MA, Hu H, Schwartz J, et al. Effect modification by vitamin D receptor genetic polymorphisms in the association between cumulative lead exposure and pulse pressure: a longitudinal study. *Environ Health* 2015;14:5.
71. Onalaja AO, Claudio L. Genetic susceptibility to lead poisoning. *Environ Health Perspect* 2000;108 Suppl 1:23-28.
72. Warrington NM, Zhu G, Dy V, et al. Genome-wide association study of blood lead shows multiple associations near ALAD. *Hum Mol Genet* 2015;24(13):3871-3879.
73. Gundacker C, Gencik M, Hengstschlager M. The relevance of the individual genetic background for the toxicokinetics of two significant neurodevelopmental toxicants: mercury and lead. *Mutat Res* 2010;705(2):130-140.
74. Kim J, Lee Y, Yang M. Environmental exposure to lead (Pb) and variations in its susceptibility. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev* 2014;32(2):159-185.
75. Zhou S, Morozova TV, Hussain YN, et al. The Genetic Basis for Variation in Sensitivity to Lead Toxicity in Drosophila melanogaster. *Environ Health Perspect* 2016;124(7):1062-1070.
76. Claussnitzer M, Dankel SN, Kim KH, et al. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N Engl J Med* 2015;373(10):895-907.
77. Sebert S, Salonurmi T, Keinanen-Kiukaanniemi S, et al. Programming effects of FTO in the development of obesity. *Acta physiologica (Oxford, England)* 2014;210(1):58-69.
78. Fitzmaurice AG, Rhodes SL, Lulla A, et al. Aldehyde dehydrogenase inhibition as a pathogenic mechanism in Parkinson disease. *Proc Natl Acad Sci U S A* 2013;110(2):636-641.

79. Fitzmaurice AG, Rhodes SL, Cockburn M, et al. Aldehyde dehydrogenase variation enhances effect of pesticides associated with Parkinson disease. *Neurology* 2014;82(5):419-426.
80. van der Valk RJ, Duijts L, Timpson NJ, et al. Fraction of exhaled nitric oxide values in childhood are associated with 17q11.2-q12 and 17q12-q21 variants. *J Allergy Clin Immunol* 2014;134(1):46-55.
81. Salam MT, Lin PC, Eckel SP, et al. Inducible Nitric Oxide Synthase Promoter Haplotypes and Residential Traffic-Related Air Pollution Jointly Influence Exhaled Nitric Oxide Level in Children. *PLoS One* 2015;10(12):e0145363.
82. Roos AB, Mori M, Gronneberg R, et al. Elevated exhaled nitric oxide in allergen-provoked asthma is associated with airway epithelial iNOS. *PLoS One* 2014;9(2):e90018.
83. Meyer UA. Pharmacogenetics - five decades of therapeutic lessons from genetic diversity. *Nature reviews Genetics* 2004;5(9):669-676.
84. Chhibber A, Kroetz DL, Tantisira KG, et al. Genomic architecture of pharmacological efficacy and adverse events. *Pharmacogenomics* 2014;15(16):2025-2048.
85. Ritchie MD. The success of pharmacogenomics in moving genetic association studies from bench to bedside: study design and implementation of precision medicine in the post-GWAS era. *Hum Genet* 2012;131(10):1615-1626.
86. Hamburg MA, Collins FS. The path to personalized medicine. *N Engl J Med* 2010;363(4):301-304.
87. Johnson JA, Gong L, Whirl-Carrillo M, et al. Clinical Pharmacogenetics Implementation Consortium Guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. *Clin Pharmacol Ther* 2011;90(4):625-629.
88. Innocenti F, Cox NJ, Dolan ME. The use of genomic information to optimize cancer chemotherapy. *Semin Oncol* 2011;38(2):186-195.
89. Locharernkul C, Shotelersuk V, Hirankarn N. Pharmacogenetic screening of carbamazepine-induced severe cutaneous allergic reactions. *J Clin Neurosci* 2011;18(10):1289-1294.
90. Bustamante CD, Burchard EG, De la Vega FM. Genomics for the world. *Nature* 2011;475(7355):163-165.
91. Chen LS, Baker TB, Piper ME, et al. Interplay of genetic risk factors (CHRNA5-CHRNA3-CHRNB4) and cessation treatments in smoking cessation success. *Am J Psychiatry* 2012;169(7):735-742.
92. Bergen AW, Javitz HS, Krasnow R, et al. Nicotinic acetylcholine receptor variation and response to smoking cessation therapies. *Pharmacogenet Genomics* 2013;23(2):94-103.
93. Chen LS, Bloom AJ, Baker TB, et al. Pharmacotherapy effects on smoking cessation vary with nicotine metabolism gene (CYP2A6). *Addiction* 2014;109(1):128-137.
94. Lerman C, Schnoll RA, Hawk LW, Jr., et al. Use of the nicotine metabolite ratio as a genetically informed biomarker of response to nicotine patch or varenicline for smoking cessation: a randomised, double-blind placebo-controlled trial. *Lancet Respir Med* 2015;3(2):131-138.

Table 1. Summary of Lessons Learned from GxE Examples

Lessons	Interaction and Phenotype
1. GXE in metabolism genotype-phenotypes are usually related to ADME (absorption, distribution, metabolism, and excretion) characteristics of targeted exposures. These are an obvious place to explore biological pathways and candidate genes.	▪ Phenylketonuria and Glucose-6-Phosphate deficiency: single metabolism genes x diet/pharmacological agents ▪ *CYP2D6/PON1/ALDH2* with pesticide exposure for Parkinson's disease ▪ *NAT2* and smoking for bladder cancer ▪ *ALDH2*2* and alcohol intake for esophageal cancer ▪ *AS3MT* and arsenic for skin lesions ▪ Pharmacogenomics examples
2. GxE Discoveries can lead to environmental interventions to prevent diseases (especially in cases where presence of both are required for outcome)	▪ Phenylketonuria and Glucose-6-Phosphate deficiency: single metabolism genes x diet/pharmacological agents ▪ *CYP2C9/VKOR1* and warfarin for anticoagulation response ▪ Nicotine metabolism genes and therapy for smoking cessation ▪ Aspirin/NSAIDs use and *MGST1/IL16* for colorectal cancer
3. Temporal considerations (birth cohorts, timing of exposure, etc.) may impact GxE findings and need to be considered	▪ *PON1* and pesticide exposure for Parkinson's disease ▪ *FTO* and physical activity for BMI
4. Quality of exposure assessment impacts detection of GxE	▪ *PON1* and pesticide exposure for Parkinson's disease ▪ *FTO* and physical activity for BMI ▪ *NOS2* and traffic pollution for respiratory symptoms
5. Scale studied can impact detection of interaction	▪ *NAT2* and smoking for bladder cancer

6. Large population sizes typically needed for GxE discovery	▪ *NAT2* and smoking for bladder cancer ▪ *FTO* and physical activity for BMI (Caucasian populations) ▪ Aspirin/NSAIDs use and *MGST1/IL16* for colorectal cancer
7. Variability in exposure distribution increases power to detect G x E and importance of investigating different ethnically and geographically diverse populations	▪ *ALDH2*2* and alcohol intake for esophageal cancer ▪ *FTO* and physical activity for BMI ▪ *AS3MT* and arsenic for skin lesions ▪ *NOS2* and traffic pollution for respiratory symptoms ▪ Carbamazepine x *HLA-B* for Stevens Johnsons Syndrome
8. No GxE Method is Universally the Most Powerful - Appropriate GxE method depends on underlying assumptions; correlations between risk factors; and true GxE model	▪ *ALDH2*2* and alcohol for esophageal squamous-cell carcinoma ▪ See Gauderman et al. companion manuscript (1)
9. Studying highly exposed populations/cohorts can provide high quality exposure assessment	▪ *FTO* and physical activity for BMI ▪ *10q24.32* x arsenic and arsenic lesions
10. Model systems and functional approaches may provide GxE insights	▪ Genetics of lead susceptibility (*Drosophila* model)
	▪ *FTO* and physical activity for BMI (human tissue samples/mouse models)
	▪ *ALDH2* and pesticides for Parkinson's disease (*ex vivo* model system)
	▪ *NOS2* and traffic pollution for respiratory symptoms (biomarker study)

Abbreviations: *ALDH2*, aldehyde dehydrogenase 2; *AS3MT*, arsenite methyltransferase; CYP, cytochrome P-450; *FTO*, fat mass- and obesity; GxE, gene-environment; *HLA-B*, human leukocyte antigen-B; NAT2, N-acetyltransferase 2; NSAIDs, non-steroidal anti-inflammatory

drugs; *NOS2*, nitric oxide synthase 2; *PON1*, paraoxonase 1; *VKORC1*, Vitamin K epoxide reductase complex subunit 1