Published in final edited form as:

Genet Epidemiol. 2016 July ; 40(5): 356–365. doi:10.1002/gepi.21967.

Review of the Gene-Environment Interaction Literature in Cancer: What do we know?

Naoko I. Simonds^{1,*}, Armen A. Ghazarian^{2,*}, Camilla B. Pimentel³, Sheri D. Schully⁴, Gary L. Ellison¹, Elizabeth M. Gillanders¹, and Leah E. Mechanic¹

¹Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

²Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD

³Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA

⁴Office of Disease Prevention, National Institutes of Health, Bethesda, MD

Abstract

Background—Risk of cancer is determined by a complex interplay of genetic and environmental factors. Although the study of gene-environment (GxE) interactions has been an active area of research, little is reported about the known findings in the literature.

Methods—To examine the state of the science in GxE research in cancer, we performed a systematic review of published literature using gene-environment or pharmacogenomic flags from two curated databases of genetic association studies, the Human Genome Epidemiology (HuGE) literature finder and Cancer Genome-Wide Association and Meta Analyses Database (CancerGAMAdb), from January 1, 2001, to January 31, 2011. A supplemental search using HuGE was conducted for articles published February 1, 2011, to April 11, 2013. A 25% sample of the supplemental publications was reviewed.

Results—A total of 3,019 articles were identified in the original search. From these articles, 243 articles were determined to be relevant based on inclusion criteria (more than 3,500 interactions). From the supplemental search (1,400 articles identified), 29 additional relevant articles (1,370 interactions) were included. The majority of publications in both searches examined GxE in colon, rectal, or colorectal cancer types; breast; or lung cancer. Specific interactions examined most frequently included environmental factors categorized as energy balance (e.g., body mass index (BMI), diet), exogenous (e.g., oral contraceptives) and endogenous hormones (e.g., menopausal status), chemical environment (e.g., grilled meats), and lifestyle (e.g., smoking, alcohol intake). In both searches, the majority of interactions examined were using loci from candidate genes studies and none of the studies were genome-wide interaction studies (GEWIS). The most commonly

Address Correspondence to: Leah E. Mechanic, Ph.D., M.P.H., Program Director, Genomic Epidemiology Branch, Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, 9609 Medical Center Drive, Room 4-E104, Rockville, MD 20850, Phone: 240-276-6847, mechanil@mail.nih.gov.

*NIS and AAG contributed equally to this work.

The authors have no conflicts of interest to declare.

reported measure was the interaction p-value, of which a sizable number of p-values were considered statistically significant (i.e., < 0.05). In addition, the magnitudes of interactions reported were modest.

Conclusion—Observations of published literature suggest that opportunity exists for increased sample size in GxE research, including GWAS identified loci in GxE studies, exploring more GWAS approaches in GxE such as GEWIS, and improving the reporting of GxE findings.

Keywords

Gene-Environment Interaction; Literature Review; Genome Wide Association Study (GWAS); Candidate Gene

INTRODUCTION

The study of gene-environment interactions (GxE) in cancer has been an active area of research for several years [Haldane 1938; Khoury, et al. 1988; Thomas 2000]. It is widely accepted that both genetic and environmental factors are associated with the etiology of cancer [Dempfle, et al. 2008; Murcray, et al. 2009; Thomas 2010; Thomas 2000] and that a complex interplay of these factors influence cancer risk [Brennan 2002; Centers for Disease Control 2000; Rothman, et al. 2001]. The study of GxE is useful for obtaining a better estimate of population-attributable risk(s); gaining a better understanding of the biological pathways/dose-response relationships; identifying individuals who may be more susceptible to cancer; understanding heterogeneity across studies; and identifying novel genes through interactions [Boffetta, et al. 2012; Hunter 2005; Thomas 2010]. Over the past decade, the field of genetic epidemiology has evolved from candidate gene and candidate gene-gene interaction or GxE studies to genome-wide association studies (GWAS). Recently, the field has begun exploring genome-wide GxE-wide interaction studies (GEWIS) [Hutter, et al. 2013; Khoury and Wacholder 2009].

The National Institutes of Health (NIH) has made the study of GxE a research priority since 2000 [Sellers 2006] as evidenced by the multitude of requests for applications and program announcements issued by several NIH institutes, such as the PAR-13-382 “Analysis of Genome-Wide Gene-Environment Interactions (R21)” [Department of Health and Human Services 2013] and PAR-11-032 “Methods and Approaches for Detection of Gene-Environment Interactions in Human Disease (R21)” [Department of Health and Human Services 2010]. The NIH has also started initiatives such as the Genes, Environment and Health Initiative (GEI) [National Human Genome Research Institute 2006] and, more recently, the Genetic Associations and Mechanisms in Oncology (GAME-ON) initiative [National Cancer Institute 2012] to further support GxE and GWAS research. Moreover, the National Cancer Institute (NCI) sponsored recent workshops underscoring the commitment to the study of GxE, such as Next Generation Analytic Tools for Large Scale Genetic Epidemiology Studies of Complex Diseases [Mechanic, et al. 2011] and the Gene-Environment Think Tank Meeting [Hutter, et al. 2013] to discuss the state of the science, identify obstacles in genetic epidemiology research, and propose solutions for epidemiological research to better understand how GxE contribute to disease.

In contrast to the hundreds of single-nucleotide polymorphisms (SNPs) that have known association with cancer [Chung and Chanock 2011; Hindorff, et al. 2009; Welter, et al. 2014], there have been much fewer statistically significant replicated GxE findings in cancer research. This is despite extensive study of well-known environmental exposures (e.g., dietary factors, smoking, hormone replacement therapy) [Aschard, et al. 2012; Hutter, et al. 2013; Kraft and Aschard 2015]. However, it should be noted that the criteria for determining “significance” for GxE findings when applying to GEWIS studies in particular have been a matter of debate [Hutter, et al. 2013].

Due to the recognized importance of GxE research in cancer, we conducted a review of the published GxE literature to gain further insight into the state of the science and to identify the presence of scientific gaps and opportunities to advance GxE research in cancer.

METHODS

Publication Search

The strategy for selecting relevant GxE publications included identifying unique citations, reviewing abstracts for relevance, and full article review for relevance verification and data abstraction. The GxE or pharmacogenomic filters from the Human Genome Epidemiology (HuGE) Literature Finder [Lin, et al. 2006] and GxE filter for the Cancer Genome-Wide Association and Meta Analyses Database (CancerGAMAdb) [Schully, et al. 2011] curated databases of genetic association studies were used to identify publications that investigated GxE interactions from January 1, 2001, to January 31, 2011, resulting in a total of 3,019 unique citations in the original literature search (primary search). To update the review, a supplemental search was performed in April 2013 using only the HuGE Literature Finder database to identify articles (from February 1, 2011 to April 11, 2013) for 1,400 additional unique citations. For the supplemental search, the CancerGAMAdb was not included because of limited relevant articles (i.e., the meta-analyses in this database frequently focused on main effects or did not include detailed information about GxE).

Abstracts were evaluated for inclusion of relevant publications. A publication was considered relevant if it met the following four inclusion criteria: (1) published in English; (2) examined a combination of genes and environment, (3) included at least 1,000 cases in the GxE interaction studied, and (4) investigated the GxE interaction as related to cancer risk. A minimum sample size of 1,000 cases was a criterion because of the large sample sizes required for GxE studies [Smith and Day 1984; Thomas 2010]. We excluded articles that only had sufficient number of cases after combining different cancer types or that examined overall cancer risk (i.e., any type of cancer). Relevant articles were considered for full article review. After review of the abstracts, 477 publications from the original search and 196 from the supplemental search were considered relevant for full article review. For the supplemental search, a 25% random sample (50 articles) of relevant articles was examined further for relevance and data abstraction. After review of the full article text, 243 articles from the primary search and 29 articles from the supplemental search were included for abstraction and analysis.

Data Abstraction

Data were abstracted for the individual interaction analyses or for each GxE combination that was described. Using a standardized template, the following data were abstracted: cancer type, environmental exposures, genetic variables, and estimates of interaction effects. Genetic variables included gene name and location (typically either rs number or chromosomal location, depending on data provided), alleles examined, study type (case-control, case-only, family), and origin of the SNP (candidate gene study, GWAS, or both). Environmental exposures were grouped into eight categories and included energy balance, lifestyle factors, exogenous hormones, endogenous hormones, chemical environment, drugs/treatments, infection and inflammation, and physical environment as described in Table 1 and previously [Ghazarian, et al. 2013]. Specific environmental terms (e.g., smoking, BMI, pesticides) were also captured. Finally, estimates of interaction were assessed with data on any reported measures of interaction, including interaction odds ratios (OR_{GxE}) defined as the OR estimate for the interaction term; joint odds ratios (OR_{GE}) defined as the OR estimate for combined effects of genes and environmental factor; and the p-value of interaction for these measures (i.e., p-value of the interaction term in a logistic model). When p-values in the publication were reported as adjusted or non-adjusted for multiple comparisons, the non-adjusted p-values were abstracted to allow for comparisons between studies.

Twenty percent of the eligible studies from the primary search and 10% from the supplemental search were reviewed for quality control. Multiple reviewers discussed discordant results and all articles that were considered not relevant, and consensus results were recorded. All specific environmental terms within environmental exposure categories were also reviewed for accuracy. Errors or discrepancies in categorization were corrected.

Data Analysis

Abstracted data were analyzed at a publication (e.g., cancer categories, origin of SNP, relevant papers) or interaction level (e.g., environmental exposure categories) as many papers examined multiple possible combinations of genetic variants and environmental exposures. Publication level analyses were thought to reflect studies while the interactions reflected the types of analyses performed and variables included. With the exception of the category “cancer types” which are presented at the publication level, all other categories (e.g., SNPs, environmental exposure) are presented at the interaction level.

Frequencies of variables were compared for relevant papers in the primary and supplemental analyses. To examine the frequency of p-values for interaction reported as $p < 0.05$, papers that reported p-values as statistically “non-significant” were considered $p = 0.05$, “significant” were considered $p < 0.05$, and other interaction p-values were categorized according to reported value. The reported p-value in the papers was to calculate the average statistically significant p-value of interaction. If the p-value was not reported, or reported as “significant”, “non-significant” or without a numerical value (i.e., if a paper reported a p-value for interaction as $p < 0.05$), the p-value for interaction was considered missing. Reported joint OR and interaction OR were used to estimate median values. Estimates of categorical frequencies (Proc Freq), averages and medians (Proc Means) were performed in SAS (version 9.3, Cary, NC).

To examine commonly reported statistically significant interactions ($p < 0.05$), we identified genes and environmental exposure category combinations for which multiple distinct publications (based on PubMed ID) reported interactions using an interaction p-value of < 0.05 . Gene names were based on names provided in the publication. However, different names may have been used to describe the same gene (i.e., *PTGS2* and *COX2*; *PTGS1* and *COX1*; *XPD* and *ERCC2*); these were categorized as the same gene. Analysis was limited to single genes and did not include those analyses where combinations of genes with environmental exposures were examined. In addition, this analysis was limited to publications where the gene name was listed in the report.

RESULTS

We conducted a review of published literature and identified 243 eligible articles in the primary search and 29 eligible articles in the supplemental search (Figure 1). From the primary search, there were 3,526 GxE interactions, and from the supplementary search, 1,370 GxE interactions. All relevant articles included in this report are listed in Supplemental Table 1.

Cancer Types

The most commonly studied cancer types were breast, lung and colorectal cancers (Figure 2), accounting for approximately 70% of all publications reviewed. All other cancer types each attributed to less than 10% of the total types studied.

Candidate and GWAS SNPs

The majority of GxE were genetic variants from candidate genes (82%, $N=2,898$ candidate gene polymorphisms; 12%, $N=416$ GWAS; 6%, $N=212$ both). The number of individual GxE examined using genetic variants identified from GWAS was greater in the supplemental search than in the primary search (13%, $N=180$ candidate gene vs. 87%, $N=1,190$ GWAS). None of these publications were GEWIS.

Environmental Exposure Categories

Figure 3A illustrates the distribution of environmental exposures in the GxE reported in relevant publications from the primary search. The most frequently studied environmental exposure was “energy balance” (41%, $N=1,457$), followed by “lifestyle” (22%, $N=760$) and “exogenous hormones” (13%, $N=456$). Within each of these three environmental exposure categories, the two most commonly studied exposures per category were smoking, alcohol (i.e., lifestyle factors), dietary factors and anthropometrics (i.e., energy balance), and hormone replacement therapy and oral contraceptive use (i.e., exogenous hormones). In the supplemental search, the most commonly examined environmental exposures were exogenous hormones (33%, $N=456$), followed by lifestyle (28%, $N=379$), endogenous hormones (21%, $N=291$), and energy balance (16%, $N=221$) (Figure 3B). In both searches, there were few interactions reported with physical environmental exposures, such as imaging or UV exposure, accounting for less than 1% of all GxE interactions.

Assessment of Interaction

Table 2 summarizes how GxE were quantified or evaluated. The most commonly reported measure was the interaction p-value. Among the interactions tested, 418 (14%) in the primary search and 113 (8.7%) in the supplemental search were statistically significant at $p < 0.05$. The average p-value = 0.017 in the primary search (from 390 p-values for interaction) and 0.023 in the supplemental search (from 112 reported p-values of interaction). Few interactions were assessed by examining OR of joint effects or measures on the additive scale.

To estimate the magnitudes of interactions based on joint odds ratio (OR_{ge}) or interaction odds ratio (OR_{gxe}), we limited to interactions with reported p-values of interaction $p < 0.05$. In the primary search, the median OR_{gxe} was 1.30 (range: 0.43–2.19; N=29 interactions) and OR_{ge} was 1.30 (range: 0.43–3.60; N=53 interactions). In the supplemental search, the median OR_{gxe} was 1.03 (range: 0.65–1.35; N=8 interactions) and median OR_{ge} was 0.99 (range: 0.33–2.17; N=71 interactions).

Most Frequently Reported Interactions

We examined any statistically significant GxE interaction reported in multiple publications. By limiting to statistically significant interaction findings (combining the primary and supplement literature review), several combinations of genes with environmental exposure categories were observed in at least two publications (Table 3). Specifically, interactions that were observed in at least three publications included: *NAT2* × lifestyle (7 publications: 4 bladder; 1 breast; 1 colon, rectal and colorectal; 1 lung; 16/16 interactions with smoking) [Ambrosone, et al. 2008; Cleary, et al. 2010; García-Closas, et al. 2005; Moore, et al. 2011; Rothman, et al. 2010; Rothman, et al. 2007; Zhou, et al. 2002], *XRCC1* × lifestyle (5 publications: 2 breast, 2 lung, 1 bladder cancer; 6/6 interactions with smoking) [Hao, et al. 2006; Pachkowski, et al. 2006; Shen, et al. 2005a; Stern, et al. 2009; Zhou, et al. 2003], *CYP1A1* × lifestyle (3 publications: 2 lung, 1 oral and pharyngeal cancer; 4/4 interactions with smoking) [Le Marchand, et al. 2003; Rotunno, et al. 2009; Varela-Lema, et al. 2008], and *MTHFR* and energy balance (3 publications: 2 breast, 1 endometrial cancer; 2/4 interactions folate plus riboflavin and 2/4 interactions folate) [Shrubsole, et al. 2004; Xu, et al. 2007b; Xu, et al. 2007c].

DISCUSSION

The present literature review of GxE studies in cancer was conducted to evaluate the focus of these studies, summarize commonly observed interactions, estimate magnitude of GxE, and identify potential research gaps. The most commonly studied cancer sites were breast, colorectal, and lung cancers. Moreover, over the 12-year observation period, we noted few consistently reported GxE findings, variability in the analytic approaches and reporting methods, frequent examination of the common cancers, and frequent use of candidate genes. In addition, a larger than expected percentage of reported p-values were < 0.05 and the magnitude of reported interactions was modest.

The results of this literature review are consistent with the state of genetic epidemiology of cancer prior to 2013, with the vast majority of studies examining GxE interactions using the candidate gene approach. Although there were proportionally more loci from GWAS explored in the supplemental search, this increase was modest, and none of the papers were GEWIS or agnostic GxE searches in GWAS. Additional GxE research opportunities may exist by exploiting the ability to look at a large number of genetic variants to study GxE interactions, and several recent studies published after our observation period performed genome-wide GxE analyses to study interactions [Du, et al. 2014; Figueroa, et al. 2014; Nan, et al. 2015; Wu, et al. 2012].

A number of statistically significant interactions were observed across studies. One of the most frequently reported statistically significant GxE finding is the increase in bladder cancer risk observed among smokers with the N-acetyltransferase 2 (*NAT2*) slow acetylation genotype [García-Closas, et al. 2005; Garcia-Closas, et al. 2013]. This association has consistently been observed since 1979, when Lower and colleagues found increased bladder cancer risk among individuals who had the slow acetylator phenotype and were exposed to aromatic amines [Lower, et al. 1979]. What makes the *NAT2* GxE interaction particularly robust is the strong biological plausibility for the finding because individuals with slow acetylation have decreased capacity to detoxify aromatic monoamines found in tobacco smoke. Consistent with this role of *NAT2* in detoxification, several reports observed association of *NAT2* genetic variants with metabolite concentrations, or ratios of specific metabolites in human blood or urine [Raffler, et al. 2015; Shin, et al. 2014; Suhre, et al. 2011a; Suhre, et al. 2011b]. Furthermore, similar associations were reported across populations and replicated in multiple studies, providing additional evidence for this interaction [García-Closas, et al. 2005; Gu, et al. 2005; Moore, et al. 2011; Yuan, et al. 2008].

Many of the publications for the most commonly reported statistically significant interactions (*NAT2* × lifestyle, *XRCC1* × lifestyle, *CYP1A1* × lifestyle, and *MTHFR* × energy balance) also detected main effects associations of these genes with the same cancer types [García-Closas, et al. 2005; Hao, et al. 2006; Le Marchand, et al. 2003; Rothman, et al. 2010; Rothman, et al. 2007; Varela-Lema, et al. 2008; Xu, et al. 2007c]. However, these main effect associations were not observed in all reports [Ambrosone, et al. 2008; Cleary, et al. 2010; Moore, et al. 2011; Pachkowski, et al. 2006; Rotunno, et al. 2009; Shen, et al. 2005a; Shrubsole, et al. 2004; Stern, et al. 2009; Xu, et al. 2007b; Zhou, et al. 2003; Zhou, et al. 2002]. In reviewing the NHGRI GWAS Catalogue using the NCBI Phenotype-Genotype Integrator (PhenGenI) [National Center for Biotechnology Information 2016] for these four genes, only *NAT2* was reported as a GWAS finding for cancer [Rothman, et al. 2010].

We also examined the approach used to estimate/report the interaction effects in GxE publications. As noted by Hutter and colleagues [Hutter, et al. 2013], most GxE interaction studies model interaction terms and scan p-values without considering full joint effects. Consistent with this observation, most articles reported a p-value for interaction, but few studies looked at combined joint effects of the genetic and environmental terms and even fewer papers examined interactions on the additive scale. Furthermore, approximately 15% of the studies in the primary search did not report any measure and described the

relationships without quantifying the interaction. Some recent studies in bladder and breast cancer suggest that considering additive effects of genetic and environmental factors may provide benefit for risk stratification [Garcia-Closas, et al. 2014; Garcia-Closas, et al. 2013], highlighting the importance of understanding the joint effects of gene and environmental exposures. As the field moves towards characterization of GxE, there will be a need for increased reporting of joint effects.

In addition to exploring approaches used to estimate interactions, this review examined the magnitude of interaction effects and levels of statistical significance. Notably, approximately 10% of the interactions that were examined were reported as a p-value of <0.05 . This large number of reported statistically significant interactions suggests potential publication bias [Ghazarian, et al. 2013; Hutter, et al. 2013]. Our observation of percentage of interactions observed with $p<0.05$ is consistent with a report from cardiometabolic traits using CardioGxE [Parnell, et al. 2014]. More importantly, the average interaction p-values in this literature review were less stringent than cut-offs typically used for main effects in GWAS ($p<10^{-8}$). However, it should be noted that the criteria for declaring significance of an observed GxE in discovery studies is unclear [Hutter, et al. 2013], although recent studies used $p<10^{-8}$ [Du, et al. 2014; Nan, et al. 2015]. Furthermore, in this review the magnitude of the joint and interactions ORs were modest. Taken together, these results suggest a need for much larger sample sizes for future studies of GxE.

We acknowledge a number of important limitations. By excluding articles reporting on less than 1,000 cancer cases, we excluded GxE studies evaluating rare cancers and may not provide a complete picture of the scope of GxE interactions reported. It was surprising that approximately 60% of papers that were excluded from the literature review was due to small sample size (i.e., less than 1,000 cases in the interaction studied) given the estimate of samples sizes for GxE as four times greater than studies of main effects [Smith and Day 1984]. Another limitation was the use of different gene names (e.g., *COX2*, *PTGS2*) or lack of gene names for GWAS findings, making it possible that other commonly observed GxE findings were missed. Finally, several papers also used $p<0.05$ as the cut-off to indicate statistical significance and many did not adjust for multiple testing, making many of the significant findings possible to be false positives and/or a product of publication bias.

However, a major strength of this study is our evaluation of a large amount of data—approaches used to estimate GxE association, information about the associations, types of genetic and environmental exposures, magnitude of interactions, and many cancer sites for a greater than 10-year period. By conducting a comprehensive review of the literature, we found opportunities in GxE research that may have been missed by conducting a review with a narrower objective. Not only is there a need to increase sample size in GxE research, but also opportunities to explore more GWAS approaches in GxE such as GEWIS.

Moreover, by abstracting a large amount of data, we found that there was a wide variability in reporting of methods and results which not only limited our ability to assess the strengths of evidence, but suggests the need for detailed and uniform reporting of GxE data. For example, it was often unclear in reports which interactions were explored and not reported and whether this was due to lack of a significant finding. Therefore, it would be helpful if

authors reported the interaction tests performed and results of these tests even if results were not statistically significant. As noted above, most articles reported a p-value for interaction, but few studies looked at combined joint effects of GxE. Furthermore, approximately 15% of the studies in the primary search did not report any measure and described the associations observed without quantifying the interaction. To better compare results across studies, at a minimum, GxE studies should consider reporting an interaction odds ratio and a p-value for interaction. This idea of providing a framework on the types of specific information that should be included in GxE reports has been provided in the STrengthening the REporting of Genetic Association studies (STREGA) initiative built on the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) [von Elm, et al. 2007] Statement by adding items—such as population stratification, Hardy-Weinberg equilibrium, selection of participants, rationale for choice of genes and variants, statistical methods—to the STROBE checklist. Although “the STREGA recommendations do not prescribe or dictate how a genetic association study should be designed,” it does encourage the transparency of reporting regardless of study design or analysis approach [Little, et al. 2009]. Applying these types of standards could facilitate interpretation of results from GxE research and the field may benefit from developing a consensus on the specific elements recommended for reporting GxE studies.

The findings from our literature review suggest some gaps and possible opportunities in GxE research. Those include broadening the spectrum of cancer types being investigated, performing more discovery using GWAS loci and GEWIS approaches, needs for larger sample sizes for these studies, and developing a more standardized method of reporting GxE methods and results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors acknowledge Mindy Clyne and Wei Yu (CDC) for assistance with HuGE database and providing the downloaded database with selected flags for this analysis. The abstracted data used for the analysis are available on request.

References

- Ambrosone CB, Kropp S, Yang J, Yao S, Shields PG, Chang-Claude J. Cigarette smoking, N-acetyltransferase 2 genotypes, and breast cancer risk: pooled analysis and meta-analysis. *Cancer Epidemiol Biomarkers Prev.* 2008; 17(1):15–26. [PubMed: 18187392]
- Aschard H, Lutz S, Maus B, Duell E, Fingerlin T, Chatterjee N, Kraft P, Van Steen K. Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. *Human Genetics.* 2012; 131(10):1591–1613. [PubMed: 22760307]
- Boffetta P, Winn DM, Ioannidis JP, Thomas DC, Little J, Smith GD, Coglian VJ, Hecht SS, Seminara D, Vineis P, 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. [PubMed: 22596931]
- Brennan P. Gene-environment interaction and aetiology of cancer: what does it mean and how can we measure it? *Carcinogenesis.* 2002; 23(3):381–7. [PubMed: 11895852]

- Centers for Disease Control. Gene-Environment Interaction Fact Sheet. Center for Disease Control and Prevention; 2000.
- Chen Y, Pei J. Factors influencing the association between CYP17 T34C polymorphism and the risk of breast cancer: meta-regression and subgroup analysis. *Breast Cancer Res Treat*. 2010; 122(2):471–81. [PubMed: 20043206]
- Chung CC, Chanock SJ. Current status of genome-wide association studies in cancer. *Hum Genet*. 2011; 130(1):59–78. [PubMed: 21678065]
- Cleary SP, Cotterchio M, Shi E, Gallinger S, Harper P. Cigarette smoking, genetic variants in carcinogen-metabolizing enzymes, and colorectal cancer risk. *Am J Epidemiol*. 2010; 172(9):1000–14. [PubMed: 20937634]
- Cui R, Kamatani Y, Takahashi A, Usami M, Hosono N, Kawaguchi T, Tsunoda T, Kamatani N, Kubo M, Nakamura Y, et al. Functional variants in ADH1B and ALDH2 coupled with alcohol and smoking synergistically enhance esophageal cancer risk. *Gastroenterology*. 2009; 137(5):1768–75. [PubMed: 19698717]
- Dempfle A, Scherag A, Hein R, Beckmann L, Chang-Claude J, Schafer H. Gene-environment interactions for complex traits: definitions, methodological requirements and challenges. *Eur J Hum Genet*. 2008; 16(10):1164–72. [PubMed: 18523454]
- Department of Health and Human Services. Methods and Approaches for Detection of Gene-Environment Interactions in Human Disease (R21). 2010. PAR -11-032<http://grants.nih.gov/grants/guide/pa-files/PA-11-032.html>
- Department of Health and Human Services. Analysis of Genome-Wide Gene-Environment (G × E) Interactions (R21). 2013. PAR-13-382<http://grants.nih.gov/grants/guide/pa-files/PA-13-382.html>
- Du M, Zhang X, Hoffmeister M, Schoen RE, Baron JA, Berndt SI, Brenner H, Carlson CS, Casey G, Chan AT, et al. No evidence of gene-calcium interactions from genome-wide analysis of colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev*. 2014; 23(12):2971–6. [PubMed: 25192705]
- Einarsdottir K, Rylander-Rudqvist T, Humphreys K, Ahlberg S, Jonasdottir G, Weiderpass E, Chia KS, Ingelman-Sundberg M, Persson I, Liu J, et al. CYP17 gene polymorphism in relation to breast cancer risk: a case-control study. *Breast Cancer Res*. 2005; 7(6):R890–6. [PubMed: 16280037]
- Figueroa JD, Han SS, Garcia-Closas M, Baris D, Jacobs EJ, Kogevinas M, Schwenn M, Malats N, Johnson A, Purdue MP, et al. Genome-wide interaction study of smoking and bladder cancer risk. *Carcinogenesis*. 2014; 35(8):1737–44. [PubMed: 24662972]
- Garcia-Closas M, Gunsoy NB, Chatterjee N. Combined associations of genetic and environmental risk factors: implications for prevention of breast cancer. *J Natl Cancer Inst*. 2014; 106(11):dju305. [PubMed: 25392194]
- García-Closas M, Malats N, Silverman D, Dosemeci M, Kogevinas M, Hein DW, Tardón A, Serra C, Carrato A, García-Closas R, 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.
- Garcia-Closas M, Rothman N, Figueroa JD, Prokunina-Olsson L, Han SS, Baris D, Jacobs EJ, Malats N, De Vivo I, Albanes D, et al. Common genetic polymorphisms modify the effect of smoking on absolute risk of bladder cancer. *Cancer Res*. 2013; 73(7):2211–20. [PubMed: 23536561]
- Gates MA, Tworoger SS, Terry KL, Titus-Ernstoff L, Rosner B, De Vivo I, Cramer DW, Hankinson SE. Talc use, variants of the GSTM1, GSTT1, and NAT2 genes, and risk of epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev*. 2008; 17(9):2436–44. [PubMed: 18768514]
- Ghazarian AA, Simonds NI, Bennett K, Pimentel CB, Ellison GL, Gillanders EM, Schully SD, Mechanic LE. A review of NCI's extramural grant portfolio: identifying opportunities for future research in genes and environment in cancer. *Cancer Epidemiol Biomarkers Prev*. 2013; 22(4):501–7. [PubMed: 23462918]
- Gu J, Liang D, Wang Y, Lu C, Wu X. Effects of N-acetyl transferase 1 and 2 polymorphisms on bladder cancer risk in Caucasians. *Mutat Res*. 2005; 581(1–2):97–104. [PubMed: 15725609]
- Haldane, J. Heredity and Politics. New York: W. W. Norton Company; 1938.
- Han J, Tranah GJ, Hankinson SE, Samson LD, Hunter DJ. Polymorphisms in O6-methylguanine DNA methyltransferase and breast cancer risk. *Pharmacogenet Genomics*. 2006; 16(7):469–74. [PubMed: 16788379]

- Hao B, Miao X, Li Y, Zhang X, Sun T, Liang G, Zhao Y, Zhou Y, Wang H, Chen X, et al. A novel T-77C polymorphism in DNA repair gene XRCC1 contributes to diminished promoter activity and increased risk of non-small cell lung cancer. *Oncogene*. 2006; 25(25):3613–20. [PubMed: 16652158]
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*. 2009; 106(23):9362–7. [PubMed: 19474294]
- Hunter DJ. Gene-environment interactions in human diseases. *Nat Rev Genet*. 2005; 6(4):287–98. [PubMed: 15803198]
- Hutter CM, Mechanic LE, Chatterjee N, Kraft P, Gillanders EM. Gene-environment interactions in cancer epidemiology: a National Cancer Institute Think Tank report. *Genet Epidemiol*. 2013; 37(7):643–57. [PubMed: 24123198]
- Khoury MJ, Adams MJ Jr, Flanders WD. An epidemiologic approach to ecogenetics. *Am J Hum Genet*. 1988; 42(1):89–95. [PubMed: 3337114]
- Khoury MJ, Wacholder S. Invited commentary: from genome-wide association studies to gene-environment-wide interaction studies—challenges and opportunities. *Am J Epidemiol*. 2009; 169(2):227–30. discussion 234–5. [PubMed: 19022826]
- Kraft P, Aschard H. Finding the missing gene-environment interactions. *Eur J Epidemiol*. 2015; 30(5):353–5. [PubMed: 26026724]
- Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, Mirabello L, Jacobs K, Wheeler W, Yeager M, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am J Hum Genet*. 2009; 85(5):679–91. [PubMed: 19836008]
- Le Marchand L, Guo C, Benhamou S, Bouchardy C, Cascorbi I, Clapper ML, Garte S, Haugen A, Ingelman-Sundberg M, Kihara M, et al. Pooled analysis of the CYP1A1 exon 7 polymorphism and lung cancer (United States). *Cancer Causes Control*. 2003; 14(4):339–46. [PubMed: 12846365]
- Lin BK, Clyne M, Walsh M, Gomez O, Yu W, Gwinn M, Khoury MJ. Tracking the epidemiology of human genes in the literature: the HuGE Published Literature database. *Am J Epidemiol*. 2006; 164(1):1–4. [PubMed: 16641305]
- Lin Z, Zhang X, Tuo J, Guo Y, Green B, Chan CC, Tan W, Huang Y, Ling W, Kadlubar FF, et al. A variant of the Cockayne syndrome B gene ERCC6 confers risk of lung cancer. *Hum Mutat*. 2008; 29(1):113–22. [PubMed: 17854076]
- Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, Khoury MJ, Cohen B, Davey-Smith G, Grimshaw J, et al. STrengthening the REporting of Genetic Association Studies (STREGA)—an extension of the STROBE statement. *Genet Epidemiol*. 2009; 33(7):581–98. [PubMed: 19278015]
- Liu Z, Li G, Wei S, Niu J, Wang LE, Sturgis EM, Wei Q. Genetic variations in TERT-CLPTM1L genes and risk of squamous cell carcinoma of the head and neck. *Carcinogenesis*. 2010; 31(11):1977–81. [PubMed: 20802237]
- Lower GM Jr, Nilsson T, Nelson CE, Wolf H, Gamsky TE, Bryan GT. N-acetyltransferase phenotype and risk in urinary bladder cancer: approaches in molecular epidemiology. Preliminary results in Sweden and Denmark. *Environ Health Perspect*. 1979; 29:71–9. [PubMed: 510245]
- MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk. Polymorphisms in the BRCA1 and ABCB1 genes modulate menopausal hormone therapy associated breast cancer risk in postmenopausal women. *Breast Cancer Res Treat*. 2010; 120(3):727–36. [PubMed: 19672706]
- Mechanic LE, Chen HS, Amos CI, Chatterjee N, Cox NJ, Divi RL, Fan R, Harris EL, Jacobs K, Kraft P, et al. Next generation analytic tools for large scale genetic epidemiology studies of complex diseases. *Genet Epidemiol*. 2011; 36(1):22–35. [PubMed: 22147673]
- Mechanic LE, Millikan RC, Player J, de Cotret AR, Winkel S, Worley K, Heard K, Heard K, Tse CK, Keku T. Polymorphisms in nucleotide excision repair genes, smoking and breast cancer in African Americans and whites: a population-based case-control study. *Carcinogenesis*. 2006; 27(7):1377–85. [PubMed: 16399771]
- Milne RL, Gaudet MM, Spurdle AB, Fasching PA, Couch FJ, Benitez J, Arias Perez JI, Zamora MP, Malats N, Dos Santos Silva I, et al. Assessing interactions between the associations of common

genetic susceptibility variants, reproductive history and body mass index with breast cancer risk in the breast cancer association consortium: a combined case-control study. *Breast Cancer Res.* 2010; 12(6):R110. [PubMed: 21194473]

- Moore LE, Baris DR, Figueroa JD, Garcia-Closas M, Karagas MR, Schwenn MR, Johnson AT, Lubin JH, Hein DW, Dagnall CL, 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):182–9. [PubMed: 21037224]
- Murcray CE, Lewinger JP, Gauderman WJ. Gene-environment interaction in genome-wide association studies. *Am J Epidemiol.* 2009; 169(2):219–26. [PubMed: 19022827]
- Murtaugh MA, Sweeney C, Ma KN, Potter JD, Caan BJ, Wolff RK, Slattery ML. Vitamin D receptor gene polymorphisms, dietary promotion of insulin resistance, and colon and rectal cancer. *Nutr Cancer.* 2006; 55(1):35–43. [PubMed: 16965239]
- Nan H, Hutter CM, Lin Y, Jacobs EJ, Ulrich CM, White E, Baron JA, Berndt SI, Brenner H, Butterbach K, et al. Association of aspirin and NSAID use with risk of colorectal cancer according to genetic variants. *JAMA.* 2015; 313(11):1133–42. [PubMed: 25781442]
- National Cancer Institute. Genetic Associations and Mechanisms in Oncology (GAME-ON): A Post-Genome Wide Association Initiative. 2012. <http://epi.grants.cancer.gov/gameon/>
- National Center for Biotechnology Information. Phenotype-Genotype Integrator (PheGenI). 2016. <http://www.ncbi.nlm.nih.gov/gap/phegeni>
- National Human Genome Research Institute. The Genes, Environment and Health Initiative (GEI). 2006. <http://www.genome.gov/19518663>
- Nickels S, Truong T, Hein R, Stevens K, Buck K, Behrens S, Eilber U, Schmidt M, Haberle L, Vrieling A, et al. Evidence of gene-environment interactions between common breast cancer susceptibility loci and established environmental risk factors. *PLoS Genet.* 2013; 9(3):e1003284. [PubMed: 23544014]
- Pachkowski BF, Winkel S, Kubota Y, Swenberg JA, Millikan RC, Nakamura J. XRCC1 genotype and breast cancer: functional studies and epidemiologic data show interactions between XRCC1 codon 280 His and smoking. *Cancer Res.* 2006; 66(5):2860–8. [PubMed: 16510609]
- Pande M, Spitz MR, Wu X, Gorlov IP, Chen WV, Amos CI. Novel genetic variants in the chromosome 5p15.33 region associate with lung cancer risk. *Carcinogenesis.* 2011; 32(10):1493–9. [PubMed: 21771723]
- Parnell LD, Blokker BA, Dashti HS, Nesbeth PD, Cooper BE, Ma Y, Lee YC, Hou R, Lai CQ, Richardson K, et al. CardioGxE, a catalog of gene-environment interactions for cardiometabolic traits. *BioData Min.* 2014; 7:21. [PubMed: 25368670]
- Raffler J, Friedrich N, Arnold M, Kacprowski T, Rueedi R, Altmaier E, Bergmann S, Budde K, Gieger C, Homuth G, et al. Genome-Wide Association Study with Targeted and Non-targeted NMR Metabolomics Identifies 15 Novel Loci of Urinary Human Metabolic Individuality. *PLoS Genet.* 2015; 11(9):e1005487. [PubMed: 26352407]
- Raimondi S, Paracchini V, Autrup H, Barros-Dios JM, Benhamou S, Boffetta P, Cote ML, Dialyna IA, Dolzan V, Filiberti R, et al. Meta- and pooled analysis of GSTT1 and lung cancer: a HuGE-GSEC review. *Am J Epidemiol.* 2006; 164(11):1027–42. [PubMed: 17000715]
- Rothman N, Garcia-Closas M, Chatterjee N, Malats N, Wu X, Figueroa JD, Real FX, Van Den Berg D, Matullo G, Baris D, et al. A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. *Nat Genet.* 2010; 42(11):978–84. [PubMed: 20972438]
- Rothman N, Garcia-Closas M, Hein DW. Commentary: Reflections on G. M. Lower and colleagues' 1979 study associating slow acetylator phenotype with urinary bladder cancer: meta-analysis, historical refinements of the hypothesis, and lessons learned. *Int J Epidemiol.* 2007; 36(1):23–8. [PubMed: 17510073]
- Rothman N, Wacholder S, Caporaso NE, Garcia-Closas M, Buetow K, Fraumeni JF Jr. The use of common genetic polymorphisms to enhance the epidemiologic study of environmental carcinogens. *Biochim Biophys Acta.* 2001; 1471(2):C1–10. [PubMed: 11342183]
- Rotunno M, Yu K, Lubin JH, Consonni D, Pesatori AC, Goldstein AM, Goldin LR, Wacholder S, Welch R, Burdette L, et al. Phase I metabolic genes and risk of lung cancer: multiple polymorphisms and mRNA expression. *PLoS One.* 2009; 4(5):e5652. [PubMed: 19479063]

- Schully SD, Yu W, McCallum V, Benedicto CB, Dong LM, Wulf A, Clyne M, Khoury MJ. Cancer GAMAdb: database of cancer genetic associations from meta-analyses and genome-wide association studies. *Eur J Hum Genet.* 2011; 19(8):928–30. [PubMed: 21487441]
- Sellers TA. The beginning of the end for the epidemiologic focus on gene-environment interactions? *Cancer Epidemiol Biomarkers Prev.* 2006; 15(6):1059–60. [PubMed: 16775161]
- Setiawan VW, Doherty JA, Shu XO, Akbari MR, Chen C, De Vivo I, Demichele A, Garcia-Closas M, Goodman MT, Haiman CA, et al. Two estrogen-related variants in CYP19A1 and endometrial cancer risk: a pooled analysis in the Epidemiology of Endometrial Cancer Consortium. *Cancer Epidemiol Biomarkers Prev.* 2009; 18(1):242–7. [PubMed: 19124504]
- Shen J, Gammon MD, Terry MB, Wang L, Wang Q, Zhang F, Teitelbaum SL, Eng SM, Sagiv SK, Gaudet MM, et al. Polymorphisms in XRCC1 modify the association between polycyclic aromatic hydrocarbon-DNA adducts, cigarette smoking, dietary antioxidants, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev.* 2005a; 14(2):336–42. [PubMed: 15734955]
- Shen J, Terry MB, Gammon MD, Gaudet MM, Teitelbaum SL, Eng SM, Sagiv SK, Neugut AI, Santella RM. MGMT genotype modulates the associations between cigarette smoking, dietary antioxidants and breast cancer risk. *Carcinogenesis.* 2005b; 26(12):2131–7. [PubMed: 16014702]
- Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, Arnold M, Erte I, Forgetta V, Yang TP, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet.* 2014; 46(6):543–50. [PubMed: 24816252]
- Shrubsole MJ, Gao YT, Cai Q, Shu XO, Dai Q, Hebert JR, Jin F, Zheng W. MTHFR polymorphisms, dietary folate intake, and breast cancer risk: results from the Shanghai Breast Cancer Study. *Cancer Epidemiol Biomarkers Prev.* 2004; 13(2):190–6. [PubMed: 14973091]
- Slattery ML, Curtin K, Poole EM, Duggan DJ, Samowitz WS, Peters U, Caan BJ, Potter JD, Ulrich CM. Genetic variation in C-reactive protein in relation to colon and rectal cancer risk and survival. *Int J Cancer.* 2011a; 128(11):2726–34. [PubMed: 20949557]
- Slattery ML, Herrick JS, Lundgreen A, Wolff RK. Genetic variation in the TGF-beta signaling pathway and colon and rectal cancer risk. *Cancer Epidemiol Biomarkers Prev.* 2011b; 20(1):57–69. [PubMed: 21068203]
- Smith PG, Day NE. The design of case-control studies: the influence of confounding and interaction effects. *Int J Epidemiol.* 1984; 13(3):356–65. [PubMed: 6386716]
- Stern MC, Lin J, Figueroa JD, Kelsey KT, Kiltie AE, Yuan JM, Matullo G, Fletcher T, Benhamou S, Taylor JA, et al. Polymorphisms in DNA repair genes, smoking, and bladder cancer risk: findings from the international consortium of bladder cancer. *Cancer Res.* 2009; 69(17):6857–64. [PubMed: 19706757]
- Suhre K, Shin SY, Petersen AK, Mohny RP, Meredith D, Wägele B, Altmaier E, Deloukas P, Erdmann J, Grundberg E, et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature.* 2011a; 477(7362):54–60. [PubMed: 21886157]
- Suhre K, Wallaschofski H, Raffler J, Friedrich N, Haring R, Michael K, Wasner C, Krebs A, Kronenberg F, Chang D, et al. A genome-wide association study of metabolic traits in human urine. *Nat Genet.* 2011b; 43(6):565–9. [PubMed: 21572414]
- Theodoratou E, Farrington SM, Tenesa A, McNeill G, Cetnarskyj R, Barnetson RA, Porteous ME, Dunlop MG, Campbell H. Modification of the inverse association between dietary vitamin D intake and colorectal cancer risk by a FokI variant supports a chemoprotective action of Vitamin D intake mediated through VDR binding. *Int J Cancer.* 2008; 123(9):2170–9. [PubMed: 18709640]
- Thomas D. Gene-environment-wide association studies: emerging approaches. *Nat Rev Genet.* 2010; 11(4):259–72. [PubMed: 20212493]
- Thomas DC. Genetic epidemiology with a capital “E”. *Genet Epidemiol.* 2000; 19(4):289–300. [PubMed: 11108640]
- Travis RC, Reeves GK, Green J, Bull D, Tipper SJ, Baker K, Beral V, Peto R, Bell J, Zelenika D, et al. Gene-environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study. *Lancet.* 2010; 375(9732):2143–51. [PubMed: 20605201]
- Truong T, Hung RJ, Amos CI, Wu X, Bickeboller H, Rosenberger A, Sauter W, Illig T, Wichmann HE, Risch A, et al. Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and

- 6p21: a pooled analysis from the International Lung Cancer Consortium. *J Natl Cancer Inst.* 2010; 102(13):959–71. [PubMed: 20548021]
- Varela-Lema L, Taioli E, Ruano-Ravina A, Barros-Dios JM, Anantharaman D, Benhamou S, Boccia S, Bhisey RA, Cadoni G, Capoluongo E, et al. Meta-analysis and pooled analysis of GSTM1 and CYP1A1 polymorphisms and oral and pharyngeal cancers: a HuGE-GSEC review. *Genet Med.* 2008; 10(6):369–84. [PubMed: 18496222]
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007; 370(9596):1453–7. [PubMed: 18064739]
- Wedren S, Lovmar L, Humphreys K, Magnusson C, Melhus H, Syvanen AC, Kindmark A, Landegren U, Fermer ML, Stiger F, et al. Oestrogen receptor alpha gene haplotype and postmenopausal breast cancer risk: a case control study. *Breast Cancer Res.* 2004; 6(4):R437–49. [PubMed: 15217512]
- Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorf L, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 2014; 42:D1001–6. Database issue. [PubMed: 24316577]
- Wen W, Cai Q, Xiang YB, Xu WH, Ruan ZX, Cheng J, Zheng W, Shu XO. The modifying effect of C-reactive protein gene polymorphisms on the association between central obesity and endometrial cancer risk. *Cancer.* 2008; 112(11):2409–16. [PubMed: 18383516]
- Wu C, Kraft P, Zhai K, Chang J, Wang Z, Li Y, Hu Z, He Z, Jia W, Abnet CC, 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):1090–7. [PubMed: 22960999]
- Xu WH, Dai Q, Xiang YB, Long JR, Ruan ZX, Cheng JR, Zheng W, Shu XO. Interaction of soy food and tea consumption with CYP19A1 genetic polymorphisms in the development of endometrial cancer. *Am J Epidemiol.* 2007a; 166(12):1420–30. [PubMed: 17827443]
- Xu WH, Shrubsole MJ, Xiang YB, Cai Q, Zhao GM, Ruan ZX, Cheng JR, Zheng W, Shu XO. Dietary folate intake, MTHFR genetic polymorphisms, and the risk of endometrial cancer among Chinese women. *Cancer Epidemiol Biomarkers Prev.* 2007b; 16(2):281–7. [PubMed: 17301261]
- Xu X, Gammon MD, Zhang H, Wetmur JG, Rao M, Teitelbaum SL, Britton JA, Neugut AI, Santella RM, Chen J. Polymorphisms of one-carbon-metabolizing genes and risk of breast cancer in a population-based study. *Carcinogenesis.* 2007c; 28(7):1504–9. [PubMed: 17372271]
- Yang SJ, Yokoyama A, Yokoyama T, Huang YC, Wu SY, Shao Y, Niu J, Wang J, Liu Y, Zhou XQ, et al. Relationship between genetic polymorphisms of ALDH2 and ADH1B and esophageal cancer risk: a meta-analysis. *World J Gastroenterol.* 2010; 16(33):4210–20. [PubMed: 20806441]
- Yuan JM, Chan KK, Coetzee GA, Castelao JE, Watson MA, Bell DA, Wang R, Yu MC. Genetic determinants in the metabolism of bladder carcinogens in relation to risk of bladder cancer. *Carcinogenesis.* 2008; 29(7):1386–93. [PubMed: 18544563]
- Zhou W, Liu G, Miller DP, Thurston SW, Xu LL, Wain JC, Lynch TJ, Su L, Christiani DC. Polymorphisms in the DNA repair genes XRCC1 and ERCC2, smoking, and lung cancer risk. *Cancer Epidemiol Biomarkers Prev.* 2003; 12(4):359–65. [PubMed: 12692111]
- Zhou W, Liu G, Thurston SW, Xu LL, Miller DP, Wain JC, Lynch TJ, Su L, Christiani DC. Genetic polymorphisms in N-acetyltransferase-2 and microsomal epoxide hydrolase, cumulative cigarette smoking, and lung cancer. *Cancer Epidemiol Biomarkers Prev.* 2002; 11(1):15–21. [PubMed: 11815396]

Figure 1.

Strategy used for GxE literature search. Primary and Supplemental literature searches and selection of relevant articles were performed as described in the “Materials and Methods”. In brief, abstracts for relevant citations were reviewed for inclusion followed by review of complete papers. Data from relevant papers was abstracted for further analysis.

Figure 2.

Distribution of the number of GxE interactions examined by cancer site in primary (A) and supplemental (B) literature searches. 3,526 interactions examined in the primary search and 1,370 interactions in the supplemental search from relevant publications.

Figure 3.

Distribution of the number of GxE interactions examined by environmental exposure category in primary (A) and supplemental (B) literature searches. 3,526 interactions examined in the primary search and 1,370 interactions in the supplemental search from relevant publications.

Table 1

Most common specific exposures identified for each environmental exposure category

Exposure Category	Most Common Specific Exposures ^I
Primary Search	
Energy Balance	Dietary Factors (e.g., specific nutrients or vitamins, vegetable intake; N=1,168); Anthropometrics (e.g., BMI, height; N=277); Physical Activity (N=12)
Lifestyle	Smoking (N=661); Alcohol (N=79); Breastfeeding (N=19)
Exogenous Hormones	Hormone Replacement Therapy (N=256); OC use (N=28); Recent Hormone Exposure (N=70)
Chemical Environment	Grilled foods/meats and heterocyclic amines (N=338); Aromatic adducts and amines (N=4)
Endogenous Hormones	Menopausal Status (N=103); Age of Menarche (N=80); Parity/Number of births (N=43)
Drugs/Treatment	NSAIDS (N=119); Statin (N=1)
Infection and Inflammation	Helicobacter Pylori (N=13); Autoimmune disease (N=11); Hay Fever (N=7); Emphysema (N=7)
Physical Environment	X-rays (N=10); Mammograms (N=8); Ultraviolet/Sun exposure (N=6)
Supplemental Search	
Energy Balance	Anthropometrics (e.g. BMI, Height; N=138); Physical Activity (N=55); Dietary Factors (e.g. specific nutrients or vitamins, vegetable intake; N=28)
Lifestyle	Smoking (N=228); Alcohol (N=82); Breastfeeding (N=69)
Exogenous Hormones	OC use (N=188); Hormone Replacement Therapy (N=158); Recent Hormone Exposure/Current Hormone Replacement Therapy use (N=110)
Endogenous Hormones	Parity/Number of Births (N=140); Age at First Birth (N=73); Age of Menarche (N=72)
Drugs/Treatment	NSAIDS (N=22)
Infection and Inflammation	Inflammation score (N=1)

BMI: body mass index, OC: oral contraceptives, NSAIDS: nonsteroidal anti-inflammatory drugs,

^I Specific exposures were grouped. For example, smoking dose (packs/day), smoking status, and smoking duration (years) were considered smoking. OC use and Hormone Replacement therapy included duration and formulations when specified by the publication.

N = number of interactions examined

Table 2

Assessment of GxE interactions by method of quantification

Method of Quantification	Number of Interactions Examined	Percent
<i>Primary Search</i>		
ORgxe	298	8.5%
ORge	283	8.0%
Additive Scale	96	2.7%
p-interaction	2,945	83.5%
Any Measure	3,050	86.5%
<i>Supplemental Search</i>		
ORgxe	1,092	79.7%
ORge	19	1.3%
Additive Scale	47	3.4%
p-interaction	1,303	95.1%
Any Measure	1,312	96.0%

ORgxe: interaction odds ratio, ORge: joint odds ratio

Table 3

Most Frequently Reported Interactions Based on Interaction P-value <0.05

Gene/Locus	Environmental Exposure Category	Number of Publications Reporting	References
<i>NAT2</i>	Lifestyle	7	[Ambrosone, et al. 2008; Cleary, et al. 2010; García-Closas, et al. 2005; Moore, et al. 2011; Rothman, et al. 2010; Rothman, et al. 2007; Zhou, et al. 2002]
<i>XRCC1</i>	Lifestyle	5	[Hao, et al. 2006; Pachkowski, et al. 2006; Shen, et al. 2005a; Stern, et al. 2009; Zhou, et al. 2003]
<i>CYP1A1</i>	Lifestyle	3	[Le Marchand, et al. 2003; Rotunno, et al. 2009; Varela-Lema, et al. 2008]
<i>MTHFR</i>	Energy Balance	3	[Shrubsole, et al. 2004; Xu, et al. 2007b; Xu, et al. 2007c]
<i>15q25</i>	Lifestyle	2	[Landi, et al. 2009; Truong, et al. 2010]
<i>ADH1B</i>	Lifestyle	2	[Cui, et al. 2009; Yang, et al. 2010]
<i>ALDH2</i>	Lifestyle	2	[Cui, et al. 2009; Yang, et al. 2010]
<i>CASP8</i>	Lifestyle	2	[Nickels, et al. 2013; Travis, et al. 2010]
<i>CRP</i>	Energy Balance	2	[Slattery, et al. 2011a; Wen, et al. 2008]
<i>CYP17A1</i>	Endogenous Hormones	2	[Chen and Pei 2010; Einarsson, et al. 2005]
<i>CYP19A1</i>	Energy Balance	2	[Setiawan, et al. 2009; Xu, et al. 2007a]
<i>ERCC6</i>	Lifestyle	2	[Lin, et al. 2008; Mechanic, et al. 2006]
<i>ESR1</i>	Energy Balance	2	[Nickels, et al. 2013; Wedren, et al. 2004]
<i>GSTT1</i>	Chemical Environment	2	[Gates, et al. 2008; Raimondi, et al. 2006]
<i>LSP1</i>	Endogenous Hormones	2	[Milne, et al. 2010; Nickels, et al. 2013]
<i>MGMT</i>	Energy Balance	2	[Han, et al. 2006; Shen, et al. 2005b]
<i>TERT-CLPTMIL</i>	Lifestyle	2	[Liu, et al. 2010; Pande, et al. 2011]
<i>TGFB1</i>	Exogenous Hormones	2	[MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk 2010; Slattery, et al. 2011b]
<i>TOX3</i>	Endogenous Hormones	2	[Milne, et al. 2010; Nickels, et al. 2013]
<i>VDR</i>	Energy Balance	2	[Murtaugh, et al. 2006; Theodoratou, et al. 2008]