Hypothesis/Commentary

Cancer Epidemiology in the 21st Century

A Review of NCI's Extramural Grant Portfolio: Identifying Opportunities for Future Research in Genes and Environment in Cancer

Armen A. Ghazarian, Naoko I. Simonds, Kelly Bennett, Camilla B. Pimentel, Gary L. Ellison, Elizabeth M. Gillanders, Sheri D. Schully, and Leah E. Mechanic

Abstract

Background: Genetic and environmental factors jointly influence cancer risk. The NIH has made the study of gene–environment (GxE) interactions a research priority since the year 2000.

Methods: To assess the current status of GxE research in cancer, we analyzed the extramural grant portfolio of the National Cancer Institute (NCI) from Fiscal Years 2007 to 2009. Publications attributed to selected grants were also evaluated.

Results: From the 1,106 research grants identified in our portfolio analysis, a random sample of 450 grants (40%) was selected for data abstraction; of these, 147 (33%) were considered relevant. The most common cancer type was breast (20%, n = 29), followed by lymphoproliferative (10%, n = 14), colorectal (9%, n = 13), melanoma/other skin (9%, n = 13), and lung/upper aerodigestive tract (8%, n = 12) cancers. The majority of grants were studies of candidate genes (68%, n = 100) compared with genome-wide association studies (GWAS) (8%, n = 12). Approximately one-third studied environmental exposures categorized as energy balance (37%, n = 54) or drugs/treatment (29%, n = 43). From the 147 relevant grants, 108 publications classified as GxE or pharmacogenomic were identified. These publications were linked to 37 of the 147 grant applications (25%).

Conclusion: The findings from our portfolio analysis suggest that GxE studies are concentrated in specific areas. There is room for investments in other aspects of GxE research, including, but not limited to developing alternative approaches to exposure assessment, broadening the spectrum of cancer types investigated, and conducting GxE within GWAS.

Impact: This portfolio analysis provides a cross-sectional review of NCI support for GxE research in cancer. Cancer Epidemiol Biomarkers Prev; 22(4); 501–7. ©2013 AACR.

Introduction

Both genes and environmental exposures have been associated with the etiology of cancer (1–4), and it is widely accepted that interactions between genetic and environmental factors influence cancer risk (5–7). Others have noted that the study of gene–environment (GxE) interactions is useful for obtaining a better estimate of population-attributable risk(s), gaining a better understanding of the relevant biologic pathways, identifying individuals who may be more susceptible to cancer, understanding heterogeneity across studies, and identifying novel genes through interactions (8–10). Advances made through the Human Genome Project and the increasingly affordable high-throughput platforms and bioinformatics capabilities have revolutionized our ability to catalog genetic variation, thereby making the search for GxE interactions more readily accessible by fostering our understanding of genetic contribution to disease. In addition, the demonstrated success of genome-wide...
association studies (GWAS) using an "agnostic" approach to interrogate the whole human genome and identify genetic variants associated with numerous diseases provide additional opportunities for studying GxE interactions using such approaches (11).

The NIH has made the study of GxE interactions a research priority since 2000 (5, 12), as evidenced by several requests for applications and program announcements issued by several NIH institutes, such as the National Cancer Institute (NCI), the National Human Genome Research Institute (NHGRI), and the National Institute of Environmental Health Sciences (NIEHS). The NIH has also started initiatives such as Genes, Environment and Health Initiative (GEI; ref. 13) which supported efforts to identify major genetic susceptibility factors for several diseases and developed technologies for reliable and reproducible measurement of environmental exposures; and more recently, the NCI funded the Genetic Associations and Mechanisms in Oncology (GAME-ON) Initiative which aims to "rapidly move forward promising leads from initial cancer GWAS by...unraveling the function of genetic variants and how environmental factors may influence the genetic effect..." in breast, prostate, ovarian, lung, and colorectal cancers (14). Moreover, NCI has sponsored recent workshops underscoring its commitment to the study of GxE interactions, such as "Next Generation Analytic Tools for Large Scale Genetic Epidemiology Studies of Complex Diseases" held on September 15–16, 2010 (15), to identify obstacles to future progress in genetic epidemiology research, including the study of GxE associations. This past year, NCI sponsored a workshop, the Gene-Environment Think Tank Meeting on January 11–12, 2012 (16), which brought together investigators in the fields of biostatistics, molecular genetics, and epidemiology to discuss the state-of-the-science, identify challenges, and propose solutions for epidemiologic research to better understand how GxE interactions contribute to disease.

Because of the recognized importance of GxE interactions, we conducted an analysis of the NCI extramural grant portfolio from Fiscal Years 2007 to 2009 to evaluate extend and focus of GxE studies, to identify potential research gaps. In addition, we examined the productivity of selected GxE grant applications by linking to research publications.

Materials and Methods

Portfolio analysis

Several databases were used to identify research grants for inclusion in the portfolio analysis including the NCI’s Portfolio Management Application as well as National Institute of Allergy and Infectious Diseases’ electronic Scientific Portfolio Assistant. In addition, the NCI Division of Extramural Activities Research Analysis and Evaluation Branch conducted a search for research grants related to GxE interactions. NCI Grants selected were those that were active for fiscal years 2007–2009, including primary projects only, research project grants, and applications that were considered human subjects applications. Grants were limited to these fiscal years with the intent of evaluating productivity from these applications in the form of research publications, that is, to allow adequate time for publication of results from the funded grant. From these applications, grants were identified that contained both a genetic and an environmental search term in the abstract and specific aims. Genetic search terms included gene(s), genetic(s), genome, genomics(s), epigenetic(s), gene mapping, DNA, RNA, protein(s), somatic, germline, single-nucleotide polymorphism (SNP), GWAS, genetic testing, and biomarker(s). Environmental search terms included social environment, social determinants, behavioral, infection, bacteria, virus, drugs, medication, personal, diet, physical activity, physical inactivity, weight, obesity, alcohol, smoking, tobacco, carcinogens, chemicals, solvents, dioxins, pesticides, metals, vinyl chloride, benzidine, diesel exhaust particles, polycyclic aromatic hydrocarbons (PAH), phthalates, polychlorinated biphenyl (PCB), phenols, bisphenol A (BPA), perfluorinated compounds, perfluoro-octanoic acid (PFOA), pytoestrogens, enterolactone (ENL), genistein, cotinine, polybrominated diphenyl ether (PBDE), organochlorine pesticides, radiation, and electromagnetic field (EMF). A total of 1,106 research grants were identified using these search criteria.

From the 1,106 research grants identified, a random sample of 450 grants (40%) was selected for data abstraction. Six individuals (E.M. Gillanders, S.D. Schully, G.L. Ellison, N.I. Simonds, A.A. Ghazarian, K. Bennett) each evaluated a sixth of those grant abstracts for relevance, that is, determination of whether grants should be considered GxE interaction applications, and if the reviewer determined the grant to be relevant, specific genetic and environmental information were abstracted from the grants. More specifically, grants were characterized using the following genetic terms: candidate gene study, GWAS, epigenetic, DNA, RNA, protein, somatic, germ-line, and biomarker. In addition, grants were characterized according to the following environmental term categories: infection and inflammation, drugs/treatment, exogenous hormones, endogenous hormones, chemical environment, physical environment, lifestyle, energy balance, methods, or general. Specific environmental terms (e.g., smoking, body mass index, and pesticides) were also captured by reviewers. Finally, data were abstracted on the cancer care continuum (primary, secondary, tertiary prevention), health disparities research (yes/no), and outcome (cancer/other).

One individual (L.E. Mechanic) reviewed a random 20% of coded grants. Any discordant results were discussed by 3 reviewers (N.I. Simonds, A.A. Ghazarian, and L.E. Mechanic) and consensus results were recorded. Additional consistency checks and review was conducted on data, such as comparison of all specific environmental terms within environmental exposure categories, review of all specific genetic terms, and cancer type compared with outcome variable (cancer/other).
**Literature review**

Relevant GxE interaction grants were linked to the NIH Spires database to evaluate the number of publications obtained from the grants. About 5,981 references linked to the 147 grants classified as GxE. However, as grant applications often have more than a single funding cycle and our goal was to evaluate the funding cycle included in this report, publications were limited to after 2007. After removing all articles published before 2007, 3,236 publications were linked to the relevant grants. As grants often have several aims, publications were limited to GxE references. To identify publications that investigated GxE interactions, we obtained a list of publications from the HuGE literature finder, or HuGENavigator (17), of articles categorized as either GxE interaction or pharmacogenomic up to January 31, 2011 (3,236 articles). Only 108 of the 3,236 articles that were linked to the GxE grant applications were also identified in the HuGENavigator search using the selected categorizations of GxE interaction or pharmacogenomic.

**Results**

The NCI grant portfolio was searched and 1,106 research grants were identified on the basis of search terms. A random sample of 450 research grants was selected for review and detailed analysis. Of the 450 grants analyzed, 147 (33%) were considered relevant based on selection criteria. Extrapolated to the larger sample of 1,106 grants, it was estimated that approximately 365 grants were examined GxE associations.

Figure 1 illustrates the different cancer types considered in the funded GxE interaction research grants. A majority of grants had cancer as their main research outcome (78%, n = 115). The most common cancer type was breast (n = 29), followed by lymphoproliferative (n = 14), colorectal (n = 13), melanoma/other skin (n = 13), and lung/upper aerodigestive tract cancers (n = 12). The cancer types with the highest amount of funding, based on fiscal year 2011, were breast (21% of funded dollars), lung (10%), prostate (9.7%), colorectal (9.0%), and leukemia (7.7%; ref. 18). Less common cancer types (<5%) included bladder, cervical, head/neck, ovarian, pancreatic, and prostate cancers. Cancers with a very low frequency among GxE interaction–funded research grants (<1%) were categorized as "other" which included brain, childhood cancers, endometrial, glioblastoma, liver, meningioma, osteosarcoma, renal, soft-tissue carcinomas, and testicular cancers. Approximately 13 grants investigated more than one cancer type (categorized as "multiple") and 8 were cancer-related but did not specify the cancer type, (these grants were cancer, not otherwise specified). Of the 13 grants categorized as multiple, 12 investigated breast cancer and at least one other cancer type. Finally, 2 grants were methods and 3 were tobacco-related cancers (i.e., cancer types that were grouped on the basis of relationship to smoking exposure).

All relevant grants were categorized for genetic and environmental exposure data using the variable definitions described in Materials and Methods. The majority of the grants were studies of candidate genes (68%, n = 100) compared with GWAS (8%, n = 12). In addition, more grants investigated germline mutations (52%, n = 76) than somatic mutations (33%, n = 51). Approximately 10% (n = 15) of grants were studies that investigated epigenetic markers.

Figure 2 illustrates the environmental terms captured in our analysis. Environmental exposures that were most frequently studied were those relating to energy...
balance ($n = 54$), drugs/treatment ($n = 43$), and lifestyle factors ($n = 33$). Other environmental categories examined included infection/inflammation, endogenous hormones, exogenous hormones, chemical environment, and physical environment. Some grants did not specify which exposures would be studied but planned to study multiple exposures. The 2 most common specific exposures within each environmental category were identified in the grant portfolio (Table 1). Diet was the most frequent energy balance exposure; non-steroidal anti-inflammatory drugs were the most frequent in the drugs/treatment category, whereas smoking was the most common lifestyle factor. Notably, the most common specific exposure, 24 of the 147 grants, was smoking.

To evaluate the productivity of GxE interaction applications, we examined the number of GxE/pharmacogenomics publications produced after 2007. From the 147 relevant grants, 108 publications were identified. These publications were linked to only 37 of the 147 grant

---

**Table 1. Most common specific exposures identified in portfolio analysis for each environmental category**

<table>
<thead>
<tr>
<th>Environmental category</th>
<th>Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection/inflammation</td>
<td>Human papillomavirus ($n = 3$)</td>
</tr>
<tr>
<td></td>
<td>Hepatitis B virus ($n = 2$), Epstein–Barr virus ($n = 2$), inflammatory</td>
</tr>
<tr>
<td></td>
<td>biomarkers ($n = 2$)</td>
</tr>
<tr>
<td>Drugs/treatment</td>
<td>Nonsteroidal anti-inflammatory drug ($n = 8$)</td>
</tr>
<tr>
<td></td>
<td>Chemotherapy drugs ($n = 4$)</td>
</tr>
<tr>
<td>Endogenous hormones</td>
<td>Endogenous hormones not otherwise specified ($n = 3$)</td>
</tr>
<tr>
<td></td>
<td>Parity ($n = 2$)</td>
</tr>
<tr>
<td>Exogenous hormones</td>
<td>Oral contraceptive use ($n = 3$), exogenous hormones not otherwise</td>
</tr>
<tr>
<td></td>
<td>specified ($n = 3$)</td>
</tr>
<tr>
<td></td>
<td>Hormone replacement therapy ($n = 2$)</td>
</tr>
<tr>
<td>Chemical environment</td>
<td>Pesticides ($n = 3$)</td>
</tr>
<tr>
<td></td>
<td>Heavy metals ($n = 2$)</td>
</tr>
<tr>
<td>Physical environment</td>
<td>Ultraviolet ($n = 10$)</td>
</tr>
<tr>
<td></td>
<td>Radiation ($n = 3$)</td>
</tr>
<tr>
<td>Lifestyle</td>
<td>Smoking ($n = 24$)</td>
</tr>
<tr>
<td>Energy balance</td>
<td>Diet ($n = 20$)</td>
</tr>
<tr>
<td></td>
<td>Alcohol use ($n = 6$)</td>
</tr>
<tr>
<td></td>
<td>Physical activity ($n = 11$)</td>
</tr>
</tbody>
</table>

**Abbreviation:** $n$, number of grants.
applications (25%). Some applications produced a large number of relevant articles with the number of articles per grant application ranging from 1 to 13 (median number of publications of 2). While most of the references were linked to only a single grant application (90%), 8 articles linked to 2 grants (7%) and 3 publications linked to 3 grants (3%).

Discussion

Studying the interaction between genetic variation and the environment to understand cancer risk has been an explicitly stated priority for the NCI. The present study conducted a cross-sectional analysis of the NCI extramural research grant portfolio of GxE interaction research to evaluate the focus of GxE studies and identify potential research gaps. The majority of identified grants and relevant articles used a candidate gene approach rather than GWAS in assessing the genetic component of GxE. Most of the environmental factors included energy balance, drugs/treatment, and lifestyle. Not surprisingly, most of the funded grants in GxE research also focused on common cancer types, notably breast cancer. These findings indicate that the focus of GxE studies have been largely concentrated in certain areas and suggest future opportunity for investments in other aspects of GxE research.

Until recently, most studies focused on examining a few candidate genes to identify possible GxE interactions. The results of this portfolio analysis are consistent with the state of genetic epidemiology of cancer when the primary focus of GWAS was to identify genetic variants associated with cancer rather than to conduct a more complex exploration of GxE interactions (15). In addition, the optimal analytical method for studying GxE on a genome-wide scale is unclear, although a few recent studies compared different approaches (19–20). Although published results from candidate gene studies resulted in few consistently replicated GxE interactions (9–11), the rapid changes in high-throughput technology, statistical methods coupled with enthusiasm to study GxE interactions in GWAS may provide additional opportunities to better examine such associations (1, 11, 21).

While genetic research has transitioned to more precise methods to better assess genetic variation, many of the frequently captured environmental exposures continue to use traditional, more qualitative methods—that is, those that can be collected through questionnaires, such as smoking history, body mass index, diet, etc. This may reflect, in part, the difficulty in conducting more sophisticated quantitative measures of exposure. The challenges in directly assessing environmental exposures may include but are not limited to the greater cost and demands placed on the participants to be measured although new technologies such as those developed by the GxE (13) and alternative study designs (e.g., 2-stage) may reduce some of these burdens (15). However, this lack of precise exposure assessment is often cited as barrier to the detection of consistent GxE interactions (15, 21). Developing more quantitative methods to record exposures is one of the areas highlighted by the NCI’s Provocative Questions Project (22), NIEHS’s 2012–2017 Strategic Plan (23), and the NIH GEI (13), suggesting that NIH appreciates the need to provide additional resources to improve exposure assessment. Some of the key research needs highlighted include technology development for exposure measurement (e.g., biological markers), new sensors and detectors, improved analytical methods, and informatics tools that can handle the large quantities of exposure data collected.

In addition to commonly studied lifestyle exposures, both over-the-counter medications and chemotherapeutic agents were frequently evaluated in the GxE grants analyzed in this study. This focus likely reflects the broad interest in personalized medicine and the desire to predict how a patient will respond to a given therapy both in terms of treatment efficacy and adverse effects (e.g., toxicity) based on genomic factors. However, most of the funded grants in the portfolio analysis focused on discovery research and little investments have been made in evaluating GxE interactions for clinical and public health application and prevention (i.e., research in evidence-based guidelines to public health impact or T2 research and beyond; ref. 24). The results from this portfolio analysis are consistent with those reported by Schully and colleagues (25) where only a small portion of the extramural grant portfolio of NCI for cancer genetics research is beyond discovery or candidate health application (e.g., test or therapy) and, as such, a greater emphasis on translational research could be warranted.

Only 25% of the relevant grants produced publications that were considered either GxE or pharmacogenomic, as assessed by linkage with literature in HuGENavigator. This result could suggest that many of the GxE interaction results from the other grant applications were null and researchers (or journals) were less interested in publishing null results or null results were more difficult to publish. Failure to publish null results could bias the GxE interaction literature, thus authors and journals should be encouraged to publish null findings. Moreover, the small proportion of applications which have published GxE interaction studies may be consistent with some of the unique challenges of doing GxE research including need for large sample sizes for adequate power, challenges in exposure assessment, problems with exposure misclassification, and the challenges of interpretation (15, 21). However, these results from the literature review analysis should be interpreted with caution because the portfolio analysis included grants for all NCI divisions which included studies other than population-based epidemiologic research, whereas the HuGE literature finder is focused on population-based analyses study designs. Therefore, it is possible that some GxE interaction articles linked to the relevant grant applications were missed.

Although the portfolio analysis suggests that investments were made to study GxE interactions in various cancer types, a quarter of funded grants investigated breast cancer. This may be reflected by a combination of...
factors. One major challenge of studying interactions is that the large sample sizes required for studies of interaction (10), where 4 times the sample size typically is needed to obtain similar statistical power than for an main effect of the same magnitude (26). Another reason for the higher investments made in breast cancer may be that well-known associations with environmental exposures exist. As such, more investigators may submit applications focusing on GxE in those areas, which may explain the difference in the number of funded grants observed between, for example, breast and prostate cancers. Also, common cancers with poor survival—such as lung—may not be studied as often (i.e., not enough study participants) as those with better survival. However, those factors do not completely explain the differences. For example, colorectal cancer, which is the third most common cancer type in both men and women (27) and has a well-known environmental component (e.g., diet), was not as commonly observed as breast cancer–funded applications.

On the basis of our portfolio analysis, very few studies were funded to develop new analytical and computational methods related to GxE research. This paucity of NIH grants funded to study GxE methods was also mentioned at the Next-Generation Analytic Tools for Large-Scale Genetic Epidemiology Studies of Complex Diseases Workshop (15). It is unclear whether this result accurately reflects the total grant portfolio considering that the selection criteria had been limited to “human subjects” which could reduce the number of methods grants identified, but the potential of this gap in research support should be examined further.

It is important to acknowledge the limitations of this portfolio analysis. First, the analysis is cross-sectional and does not reflect all of the cancer research devoted to GxE. Nevertheless, we believe this analysis provides an overall picture of the scope of NCI support for GxE research.

The findings from our portfolio analysis suggest some gaps in currently funded research in GxE and identified possible opportunities for further investments and research in GxE. Those include developing alternative approaches to exposure assessment, broadening the spectrum of cancer types being investigated, and conducting GxE within GWAS.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: E.M. Gillanders, S.D. Schully, L.E. Mechanic
Development of methodology: K. Bennett, E.M. Gillanders, S.D. Schully, L.E. Mechanic
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A.A. Ghazarian, K. Bennett, S.D. Schully, L.E. Mechanic
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A.A. Ghazarian, N.I. Simonds, K. Bennett, C.B. Pimentel, G.L. Ellison, E.M. Gillanders
Writing, review, and/or revision of the manuscript: A.A. Ghazarian, N.I. Simonds, C.B. Pimentel, G.L. Ellison, E.M. Gillanders, S.D. Schully, L.E. Mechanic
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A.A. Ghazarian, E.M. Gillanders
Study supervision: L.E. Mechanic

Acknowledgments
The authors thank Scott Rogers (NCI DCCPS, EGRP) for all of his assistance with the analysis of the grant portfolio and Mindy Clyne and Wei Yu (CDC) for assistance with HuGE database and providing the downloaded database with selected flags for this analysis.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 11, 2013; accepted February 13, 2013; published OnlineFirst March 5, 2013.

References

A Review of NCI’s Extramural Grant Portfolio: Identifying Opportunities for Future Research in Genes and Environment in Cancer

Armen A. Ghazarian, Naoko I. Simonds, Kelly Bennett, et al.


Access the most recent version of this article at: doi:10.1158/1055-9965.EPI-13-0156

This article cites 18 articles, 8 of which you can access for free at: http://cebp.aacrjournals.org/content/22/4/501.full.html#ref-list-1

This article has been cited by 1 HighWire-hosted articles. Access the articles at: /content/22/4/501.full.html#related-urls

Sign up to receive free email-alerts related to this article or journal.

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.