Investments in Cancer Genomics: Who Benefits and Who Decides

Morris W. Foster, PhD, John J. Mulvihill, MD, and Richard R. Sharp, PhD

The Cancer Genome Atlas—formerly the Human Cancer Genome Project—provides an opportunity for considering how social concerns about resource allocation are interrelated with practical decisions about specific research strategies—part of a continuing convergence between scientific and public evaluations of priorities for biomedical research funding. For example, the manner, order, and extent that The Cancer Genome Atlas selects tumor types and populations to be sampled will determine who benefits most from its findings. Those choices will be determined on the basis of both scientific and social values.

By soliciting public involvement and conducting rigorous policy analysis in the design of large scientific projects such as The Cancer Genome Atlas, cancer researchers can help democratize the allocation of scientific resources and foster public confidence in biomedical research. (*Am J Public Health.* 2006;96:1960–1964. doi:10.2105/AJPH.2005.075424)

LARGE-SCALE GENOMIC research is in the process of moving from basic science projects that provide common resources for many types of genetic research—such as the Human Genome Project, the Haplotype Map Project, and the SNP (from "single-nucleotide polymorphism") Consortium—to disease-specific projects further down the translational path. Critiques of large investments in those basic scientific resources have focused on the *relative* value of genomic research compared with alternative nongenetic strategies for allocating public funds for biomedical research.[1,2] The emerging translational focus will raise additional questions about the relative values of focusing genomic investigations on particular diseases, phenotypes, and populations.

Those latter questions are especially relevant for initiatives that coordinate significant financial resources across multiple research centers to systematically investigate disease-specific domains, and they are likely to have considerable influence in setting downstream research and clinical agendas. One such initiative is The Cancer Genome Atlas (TCGA), formerly known as the Human Cancer Genome Project.

The Human Cancer Genome Project was proposed by the National Cancer Institute (NCI) and the National Human Genome Research Institute as a $1.5 billion, 10-year project.[3] The goal of TCGA is to identify genomic alterations associated with common cancer types.[4] To accomplish this goal, the initial plan of the NCI working group that came up with the idea for the Human Cancer Genome Project was to characterize 250 samples of the 50 most common tumor types for regions of genomic loss or amplification, chromosomal rearrangements, regions of aberrant methylation, and mutations in the coding regions of all human genes. Both somatic (induced) and germline (inherited) alterations in specific genes will be examined with a variety of information, including sequence data, gene expression data, and copy number and gene loss assays.[5] The aim is to provide a comprehensive frame of reference for examining molecular processes associated with carcinogenesis, which could have significant benefits for the identification of new drug targets and the use of existing pharmacological agents. A 3-year, $100 million pilot project cofunded by the NCI and the National Human Genome Research Institute to test the feasibility and utility of a larger TCGA was announced in December 2005; initial requests for applications for work on the project were issued in June 2006.[6]

Although it was proposed as a US-only project, TCGA—like its basic science predecessors—has ready potential for expanding into an international collaborative. For example, a smaller-scale cancer genome project has been pursued in Great Britain by the Wellcome Trust Sanger Institute since 2000, but it has not sequenced the whole genomes of tumor cells and has instead focused on selected genes.[7] Collaboration discussions have occurred between leaders of the Sanger and National Institutes of Health projects.[5] Even without international collaboration, TCGA design choices will have consequences for how scientists study cancer worldwide and how benefits are distributed both inside and outside the United States.

The distribution of benefits from disease-targeted genomic projects such as TCGA will depend on which specific clinical phenotypes are investigated, the populations from which samples of those phenotypes are obtained, and the distribution of genetic alterations that contribute to disease phenotypes within and between those populations. When different choices in research design result in a greater or lesser benefit for particular groups of patients or certain populations, researchers should be prepared for TCGA and similar second-generation genomic initiatives to be subjected to considerable scrutiny by both the scientific community and the public.

A full version of TCGA may never be implemented, of course, and it may have to be scaled down in size or scope. Nonetheless, by examining its potential implications for the distribution of benefits, we can gain insight into the ethical choices and policy debates that lie ahead as genomic research enters a new phase in its development.

GENOMIC RESEARCH AND "POLITICAL SCIENCE"

Questions about resource allocation for different phenotypes and populations often are addressed by funding agencies and are isolated from questions about the design of any specific project. Those who provide support to research make decisions about the allocation of funds in accordance with their stated institutional aims and values, including considering how potential benefits may be distributed among those in need. Scientists then submit individual proposals that respond to—but do not structure—those allocation decisions. Although interested citizens have lobbied in the past for increased spending in an area of research, they have rarely lobbied on behalf of a particular research project.

AIDS patients and activists, however, have made the question of benefit distribution a public as well as scientific issue in the funding of specific biomedical studies.[8] Subsequently, patients and advocates for a number of diseases (e.g., breast cancer, pancreatic cancer, and prostate cancer) have increasingly raised ethical and policy issues in the context of scientific evaluation of research projects.[9–12] Lay representatives of such groups now regularly serve on scientific review panels and funding councils.[13]

Large disease-targeted research initiatives such as TCGA may continue to blur traditional lines between public and scientific assessments of research priorities not only because of the initiative's prominence or expense but also because of its systematic plan. TCGA embeds questions of resource allocation in its basic scientific design because of the systematic scope of the proposed effort and the high level of funding required. Already concerns have been expressed that TCGA may crowd out smaller-scale investigator-initiated cancer genetics projects that have significant overlap at specific tumor sites or that compete for the same National Institutes of Health cancer research funds.[5,6]

In addition to affecting what other kinds of cancer genetics studies may be funded, it is likely that findings from TCGA will have a major influence on cancer research more generally through the production of highly detailed information on particular tumor types. This secondary effect may determine which patients benefit most from the project's investments in cancer genomics. Thus, TCGA could be simultaneously a research design and a resource allocation plan for the future of genomic investigations of cancer, which will open TCGA to a broader range of scientific and public questions than is typical of individual cancer research projects.

WHAT SAMPLES SHOULD BE COLLECTED?

The aim of TCGA is to identify common genetic alterations that occur in at least 5% of a given tumor type.[4] Rather than search for specific genetic alterations, the project will search for particular genes with alterations at or above a threshold level greater than normal background variation. The working hypothesis is that genes with higher levels of alterations are more likely to be involved in a biological pathway through which common forms of cancer may develop. Consequently, although particular alterations may be rare, the genes and pathways that are implicated may not be. But will a focus on the common genes and pathways involved in the development of cancer result in common benefits that are equitably distributed among those affected by each tumor type? This is similar to the question about whether common diseases have common variants that contribute to them (the CD/CV hypothesis), which was a central assumption and critique of the Haplotype Map Project.[14]

Although some genes and biological pathways contribute to multiple tumor types, choices about which tumor types and subtypes to sequence will result in some cancers that have extensive, publicly available sequence information for further analysis and study and other cancers that will not have such data. The use of a "most common" criterion for selecting tumor types to investigate—whether selected on the basis of US or international incidence rates—will result in those rarer cancers not in a top-50 or top-25 list (wherever the funding line is drawn) being excluded from the direct benefits offered by TCGA.

By contrast, patient advocates often claim that because so little is known about many rare diseases, these diseases should have priority in a rational resource allocation scheme.[15] Advocates for increased funding of rarer cancers likely will be concerned about cancers that do not make the priority list and thus become funding orphans. This omission would not just be a 1-time loss from the point of view of these advocates. For those tumor types that it does characterize, TCGA will enable multiple secondary analyses by researchers worldwide through a policy of rapid release of community resource data, which could result in a knowledge and benefit gap for those types and subtypes not characterized.

At the same time, proposing to sequence the same number of samples for each phenotype on a priority list may not provide sufficient numbers of samples for the investigation of histological and molecular subtypes within the most common cancers that have the greatest influence on cancer mortality. The Sanger cancer genome project sequenced a limited number of breast tumors and identified only a few somatic alterations, which suggests that the number of samples of a particular type or subtype may be crucial to success.[16] With an equal number of samples allotted to each tumor type on TCGA's priority list, more resources will be spent proportionately on less common tumor types than on more common types when measured by incidence and mortality. Alternatively, making finer distinctions among tumor subtypes, such as allotting additional numbers of samples to each of the most common histological subtypes of breast cancer, may increase the likelihood of direct benefits for those phenotypes studied, but finer distinctions also will reduce the number of rarer tumor types and subtypes studied.

Decisions about which populations to sample in TCGA could lead to other concerns about the distribution of benefits from the project. Projects funded by the NCI often assemble a diverse set of samples that are purposefully representative of US racial/ethnic groups. An even broader policy

of inclusion might select donors whose ancestries approximate global human genetic diversity on the basis of divergent population histories, such as what was done in the International Haplotype Map Project.[17] Such sampling strategies enable the distribution of benefits among those with different inherited susceptibilities to cancers that vary in frequency according to ancestry. For example, polymorphisms in xenobiotic metabolism genes, such as the CYP2A6 gene, vary in frequency between populations and may contribute to susceptibility for lung cancer.[18] Although knowing how a gene functions is important, information about specific alterations to that gene that vary in frequency by population may be valuable for developing downstream diagnostics and preventive agents that block a particular alteration from contributing to the development of cancer. Thus, focusing on genes and pathways common to the development of cancer may not necessarily result in benefits common to all populations, which makes the sampling design a key factor in the assessment of the potential distribution of TCGA benefits.

Somatic alterations that are the result of specific environmental exposures also may vary in frequency by population.[19] However, the current working proposal for TCGA does not consider selecting tumor samples on the basis of information about donors' exposures, such as what might be inferred from detailed demographic and life history information.[4] Inferences about specific exposures, such as smoking, diet, or occupational exposures to toxins (e.g., radiation) are impossible if all we know about a tumor sample is the self-reported racial/ethnic identity of the donor. To increase the likelihood of identifying potential environmental contributors to somatic mutations associated with particular cancer types, TCGA samples could purposefully be collected to reflect a range of ages, medical histories, ages at tumor onset, social circumstances, exposure histories, occupations, geographic locations, and other characteristics better suited for serving as proxies for different exposures to environmental toxicants and carcinogens.

For example, sampling tumors from colorectal cancer patients who smoke would enrich the sample for somatic alterations associated with tobacco use.[20] Recruiting colorectal cancer patients who drink alcohol or consume meat on a daily basis would enrich the sample for alterations that may be associated with those risk factors.[21] Recruiting tumor donors with histories of ulcerative colitis or Crohn's colitis would enrich the sample for alterations that may be associated with inflammatory bowel risk factors.[22] But which of these potential associations might be used?

The proposed plan for TCGA avoids making these difficult choices by making the underlying assumption that most or all inherited and environmentally induced alterations will affect a smaller number of the same common genes and pathways. If true, TCGA can study nearly any sample of a given tumor type and still benefit many patients and populations with that type by relying on relatively simple selection criteria of disease incidence, equal numbers of samples, and a racially/ethnically representative sample of the US population. These criteria, though, do not take into account the diverse scientific and social contexts within which different tumor types are embedded, which may suggest a more complex distribution of benefits among cancer types, subtypes, populations, ancestries, and types of exposures. Because of that potential for complexity, the limited numbers of samples that could be studied by TCGA, and the magnitude of resources that would be committed to the project, mean that choices in its scientific design will necessarily involve complex social considerations.

WEIGHING DIFFERENT VALUES

Competing scientific and social values can be reconciled in several different ways.

Greater Common Good

Making choices in accordance with the likely benefit to the most people would calculate the total aggregate number of tumors globally and recruit participants from populations that best approximate global human genetic variation and those indicators that previous research has suggested contribute to the largest numbers of cancer cases worldwide. This could be modified to define *common good* as being within the US population only (the approach favored in the initial Human Cancer Genome Project proposal), although that raises the question of whether the ethical benefits of the project should be evaluated solely among those citizens whose tax dollars support the project or more globally, particularly because the distribution of benefits (or lack thereof) from TCGA will not stop at the US borders. The advantage of a greater common good criterion is that all choices in research design could be made with cost-effectiveness calculations. A consequence of this approach is an emphasis on certain kinds of contributors to cancer to the exclusion of other kinds; for example, an emphasis on environmentally induced alterations rather than inherited susceptibility to most tumor types.

Distributive Justice

The greater common good model also could be modified to maximize benefit to those populations or subpopulations that are disproportionately burdened by disease or those presently underserved by alternative investments in biomedical research on the basis of a principle of distributive justice.[23] With that principle, any allocation scheme should be guided by a commitment to produce the greatest benefit for those who are least advantaged. This does not mean that those who are already socially advantaged need to be constrained in their benefits. Rather, it implies that the least advantaged need to benefit more via that allocation strategy than any competing alternative. If extrapolated to prioritize vulnerable populations, such as children, this principle would mark those groups as having a greater moral claim on benefits. For example, the concept of *person-years-of-life saved* shows that the cure for or prevention of cancer among the very young yields decades of life to society compared with saving the life of someone who is advanced in years. Thus, although pediatric tumors are relatively low on any "most common" list of tumor types, an emphasis on these tumors might be defended by appeal to this principle. The

problem, of course, is that multiple claims on the basis of distributive justice or moral priority could compete with one another for limited research resources.

Greater Scientific Benefit

This approach would base design decisions on the potential for advancing scientific knowledge. Greater potential would not necessarily coincide with the greater common good or a principle of distributive justice, but it might include some phenotypes that are not among the most common types or subtypes, and it might overrepresent some populations or ancestries and indicators of particular environmental exposures. Nonetheless, maximizing scientific benefits in areas that are most ripe for advancement could have the effect of discovering biological mechanisms or drug targets that could be applied to other tumor types, indicators, and populations. Another way to think about benefit is in terms of potential for translational benefit, which may not be the same as potential for basic scientific benefit. The difficulty with either a greater basic or translational impact model is that the evaluation of potential for benefit inevitably is somewhat subjective, and planners would be left to choose between equally promising designs that would benefit different populations.

Eclectic Approach

Some elements of each of the previous models could be combined, particularly because the larger project will involve multiple choices for each tumor type. The advantage of an eclectic approach is that it takes into account different degrees and types of scientific knowledge and different social issues for each cancer phenotype, which may warrant the application of different approaches for setting research priorities. For example, an especially strong case might be made for the greater impact of sequencing a rare but scientifically promising phenotype, and the greater common good of sequencing a larger number of lung or colorectal tumors could be justified on the basis of the greater contribution to cancer mortality from those tumor sites. Similarly, a case might be made for enriching samples of some cancer types among those with lifetime exposures to fertilizer runoff, a contributor more likely to be found in less developed agricultural countries. The difficulty with an eclectic approach is that it might undermine the systematic advantages of a large, centrally coordinated project and thereby limit potential synergies made possible by coordinated planning and potentially make choices more vulnerable to ad hoc manipulation for achieving a preferred outcome.

A DEMOCRATIC APPROACH

Similar to scientific questions, we believe policy issues are best resolved through open discussion. Somehow, the public should be involved in assessing the potential distribution of benefits from biomedical research, a view that has increasing support on the basis of AIDS, cancer, and other patient advocacy groups efforts. By opening internal methodological discussions to include a broader range of stakeholders, the setting of research priorities can more fully embody democratic commitments and ethical ideals in research.[24] Far too often, social aspects of science are left unstated and are not given the level of critical attention they deserve. For example, although TCGA has a working group for ethics, law, and policy issues, the charge to this group is narrow and focuses primarily on informed consent, data release, and intellectual property issues that are crucial for facilitating the science of the project but do not address broader social implications of the design, such as the distribution of benefits among different groups and populations.[25] Although trade-offs between scientific and social values are inevitable in any large-scale scientific project, a lack of explicit analysis of the latter may result in decisions that fail to fully take into account relevant ethical choices. Moreover, making those choices explicit during the planning and pilot phases is critical to building public trust in the research enterprise. Unlike the International Haplotype Map Project,[17] TCGA largely has avoided any explicit public engagement about the intersection of social and scientific issues.

The challenge, of course, is identifying a process through which meaningful public involvement can be incorporated into scientific planning. On the one hand, that process must involve more than the mere symbolic inclusion of lay persons or representatives of disease advocacy groups, with 1 goal being to reach beyond activists and advocates with specific interests to sample the general public's views on how the benefits of research should be distributed. On the other hand, public participation in research planning should not create significant administrative burdens or needless research delays.

One strategy for democratizing the planning process is to catalogue the pros and cons of available study designs and to use methods from the social sciences to investigate how members of the public evaluate those options, which was done at the 4 community sites where the Haplotype Map Project samples were obtained.[17] The point of such an approach would not be to conduct a survey or poll that identifies the design most preferred by the public but to understand how nonscientists work through the same choices that project planners face. This approach would provide a context for better understanding competing ethical principles that might be applied to investing scarce resources in large-scale biomedical projects such as TCGA. Investigating those principles is different from asking the public which cancers to sequence or which populations to sample.

An advantage of this approach is that it limits the introduction of personal interests in particular cancers or populations and instead raises the more general issue of how to most fairly allocate scarce biomedical research resources with respect to the benefits that might result. Findings from such an investigation, which could be conducted parallel with the pilot phase of TCGA, would provide project planners with a better sense of how a variety of stakeholders weigh social values associated with different research designs and sampling strategies. It is likely that a variety of public views on those social considerations will emerge. Although this may or may not lead to broader public debate, it will enable project organizers to make more informed choices that include both scientific and social considerations.[24]

Such an approach also may provide an empirical basis for applying different principles in the cases of different cancer types. Instead of attempting to reach an a priori conclusion about which principle is morally superior to all others and should be applied in all cases, an empirical socially scientific investigation of the implications of applying particular principles could provide public policy rationales for making multiple sophisticated choices between competing scientific designs. ■

About the Authors

Morris W. Foster is with the Center for Applied Social Research, the Cancer Institute, and the General Clinical Research Center, University of Oklahoma, Norman. John J. Mulvihill is with the Department of Pediatrics, University of Oklahoma Health Sciences Center, Norman. Richard R. Sharp is with the Center for Medical Ethics and Health Policy, Baylor College of Medicine, Houston, Tex.

Requests for reprints should be sent to Morris W. Foster, PhD, Dept of Anthropology, University of Oklahoma, 455 W Lindsey, Rm 505C, Norman, OK 73019 (e-mail: morris.w.foster–1@ou.edu).

This article was accepted December 18, 2005.

Contributors

M.W. Foster originated the study and wrote the paper. J.J. Mulvihill contributed to the analysis and assisted with the writing. R.R. Sharp contributed to the ethical analysis and assisted with the writing.

Acknowledgments

This study was funded by the National Cancer Institute (grant CA114626–01), the National Institute on Environmental Health Sciences (grant ES11174), and the National Human Genome Research Institute (grants HG02691, HG03042 and HG03083).

Note. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the funders.

Human Participant Protection

No protocol approval was needed for this study.

References

1. Sze J, Prakash S. Human genetics, environment, and communities of color: ethical and social implications. *Environ Health Perspect.* 2004;112:740–745.

2. National Cancer Institute/National Human Genome Research Institute. The Cancer Genome Atlas. Available at: http://cancergenome.nih.gov/index.asp. Accessed July 24, 2006.

3. Kaiser J. NCI gears up for cancer genome project. *Science.* 2005;307: 1182.

4. NCAB Working Group on Biomedical Technology. Report to National Cancer Advisory Board. Bethesda, Md: National Cancer Institute; 2005.

5. Bonetta L. Going on a cancer hunt. *Cell.* 2005;123:735–737.

6. Garber K. Human Cancer Genome Project moving forward despite some doubts in community. *J Natl Cancer Inst.* 2005;97:1322–1324.

7. Wellcome Trust Sanger Institute. The Cancer Genome Project. Available at: http://www.sanger.ac.uk/genetics/CGP. Accessed July 24, 2006.

8. Resnik D. Setting biomedical research priorities: justice, science, and public participation. *Kennedy Inst Ethics J.* 2001;11:181–204.

9. Fintor L. Advocacy group update—setting the cancer agenda. *J Natl Cancer Inst.* 1993;85:607–608.

10. Waller M, Batt S. Advocacy groups for breast-cancer patients. *Can Med Assoc J.* 1995;152:829–833.

11. Young-McCaughan S, Rich IM, Lindsay GC, Bertram KA. The Department of Defense Congressionally Directed Medical Research Program: innovations in the federal funding of biomedical research. *Clin Cancer Res.* 2002;8:957–962.

12. Fernandez-Zapico ME, Kaczynski JA, Urrutia R. Pancreatic cancer research: challenges, opportunities, and recent developments. *Curr Opin Gastroenterol.* 2002;18:563–567.

13. Andejeski Y, Breslau ES, Hart E, et al. Benefits and drawbacks of including consumer reviewers in the scientific merit review of breast cancer research. *J Womens Health Gender-Based Med.* 2002;11:119–136.

14. Foster MW, Sharp RR. Will investments in biobanks, prospective cohorts, and markers of common patterns of variation benefit other populations for drug response and disease susceptibility gene discovery? *Pharmacogenomics J.* 2005;5:75–80.

15. Mishra R. New funding spurs battle against pancreatic cancer. *Boston Globe.* May 23, 2003:A1.

16. Stephens P, Edkins S, Davies H, et al. A screen of the complete protein kinase gene family identifies diverse patterns of somatic mutations in human breast cancer. *Natl Genet.* 2005;37: 590–592.

17. International HapMap Consortium. Integrating ethics and science in the International HapMap Project. *Nat Rev Genet.* 2004;5:467–475.

18. Oscarson M. Genetic polymorphisms in the cytochrome P450 2A6 (CYP2A6) gene: implications for interindividual differences in nicotine metabolism. *Drug Metab Dispos.* 2001; 29:91–95.

19. Thomas CB, Nelson DO, Pleshanov P, et al. Elevated frequencies of hypoxanthine phosphoribosyltransferase lymphocyte mutants are detected in Russian liquidators 6 to 10 years after exposure to radiation from the Chernobyl nuclear power plant accident. *Mutat Res.* 1999;439:105–119.

20. Giovannucci E. An updated review of the epidemiological evidence that cigarette smoking increases risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev.* 2001;10:725–731.

21. Potter JD. Colorectal cancer: molecules and populations. *J Natl Cancer Inst.* 1999;91:916–932.

22. Eaden J. Review article: colorectal carcinoma and inflammatory bowel disease. *Aliment Pharmacol Ther.* 2004; 20(suppl 4):24–30.

23. Rawls J. *A Theory of Justice.* Cambridge, Mass: Belknap Press of Harvard University Press; 1999.

24. Kitcher P. Responsible biology. *Bioscience.* 2004;54:331–336.

25. National Cancer Institute/National Human Genome Research Institute. The Cancer Genome Atlas: Ethics, Law, and Policy. Available at: http://cancergenome.nih.gov/components/elp.asp. Accessed July 24, 2006.