11. M. Angastiniotis, B. Modell, *Ann. N.Y. Acad. Sci.* **850**, 251 (1998).
12. E. Harris, M. Tanner, *Br. Med. J.* **321**, 817 (2000).
13. E. Schaeffeler *et al*., *Lancet* **358**, 383 (2001).
14. A. Djimdé, O. K. Doumbo, R. W. Steketee, C. V. Plowe, *Lancet* **358**, 890 (2001).
15. D. J. Weatherall, J. B. Clegg, *Genes Immun.* **3**, 331 (2002).
16. "Macroeconomics and health: Investing in health for economic development" (Report of the Commission on Macroeconomics and Health, World Health Organization, Geneva, 2001).
17. G. T. Keusch, C. A. Medlin, *Nature* **422**, 561 (2003).
18. T. Pang, *Lancet* **361**, 716 (2003).
19. This essay was based on experience as lead writer for a WHO report (*7*). I thank L. Rose for typing the essay and the Leverhulme Trust for support.

VIEWPOINT

Genomic Priorities and Public Health

Kathleen Ries Merikangas[1]* and Neil Risch[2,3]

Given the continuing difficulty of identifying genes for complex disorders in a robust, replicable manner, and the extensive resources devoted to this effort, it is becoming increasingly important to analyze the relative benefits of genomics research for public health applications and for the understanding of disease pathogenesis. To establish priorities for genetics research, we review and evaluate several characteristics of selected exemplary complex diseases, including phenotypic accuracy, knowledge of specific and nonspecific genetic and environmental risk factors, and population prevalence and impact. We propose that complex diseases with the strongest evidence for genetic etiology, limited ability to modify exposure or risk factors, and high public health impact should have the highest priority for genetics research.

Advances in molecular genetics have generated substantial progress in identifying the genetic basis of Mendelian diseases; however, the pace of the discovery of genes for complex disorders has been slower (*1*). The limited progress has generated substantial debate regarding optimal strategies and priorities for genetic studies of complex human disorders, particularly in light of the copious resources devoted to identifying susceptibility genes. Although the importance and relevance to human biology of identifying and characterizing genes for complex disorders has been amply demonstrated, it is unlikely that the current level of resources devoted to basic and applied genetic research can be sustained indefinitely. With the continuing difficulty of identifying genes for complex disorders in a robust, replicable manner, questions regarding the cost and potential applications of this work to public health will become increasingly salient (*2*, *3*). In Table 1, we describe disorder characteristics that may be evaluated in choosing appropriate strategies for genetic research. Rather than providing a comprehensive list, specific disorders were selected to illustrate variability in the key areas of phenotypic accuracy, knowledge regarding specific and nonspecific genetic and environmental risk factors, and population prevalence and impact.

How Well Can We Define the Phenotype?

The reliability and validity of phenotypic characterization for complex diseases is highly variable. Whereas tumor biopsies, insulin levels, glucose reactivity, and immunoreactivity can be measured reliably, other diseases in the table (such as autism and schizophrenia) are based primarily on diagnostic criteria obtained from a clinical interview. Our limited ability to measure and characterize these phenotypes is often the rate-limiting step in etiologic research. Research designed to identify endophenotypes [phenotypic traits or markers that may represent intermediate forms of expression between the output of underlying genes and the broader disease phenotype (*4*)] may lead to more rapid success in identifying susceptibility genes for these disorders.

What Is Known Regarding the Etiologic Role of Genes?

Familial recurrence risks, measured by λ (i.e., the ratio of the risk of disease in relatives of affected cases to the population prevalence or to relatives of controls, based on controlled family studies of first degree relatives) are shown in the first column of Table 1 under the heading "Genes" (*5*). Successful identification of genes for several complex disorders resulted from the discrimination of disease subtypes based on clustering within families (such as type 1 versus type 2 diabetes and early versus late onset of breast cancer and Alzheimer's disease). Although the absolute risk of disease in first degree relatives of cases may be low, elevated risk with respect to the population prevalence (as measured by λ) can indicate the importance of genetic susceptibility. Genetic epidemiologic studies can provide information on the extent to which the familial recurrence risk is attributable to genes. For example, although the absolute risk to relatives of multiple sclerosis cases is only 4%, it is still 20 times as great as the general population risk. Twin, adoption, and half-sibling studies have shown that the familial clus-

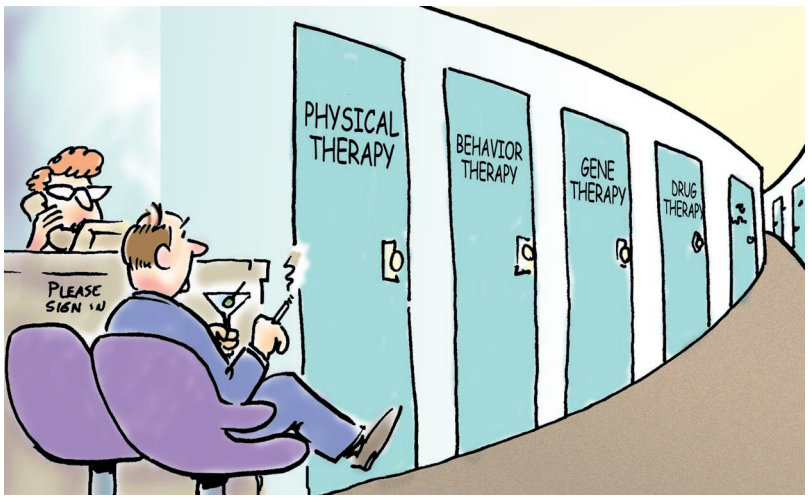

ILLUSTRATION: JOE SUTLIFF

[1]Section on Developmental Genetic Epidemiology, Mood and Anxiety Disorders Program, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, 15K North Drive, Mail Stop Code 2670, Bethesda, MD 20892, USA. [2]Department of Genetics, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305–5120, USA. [3]Division of Research, Kaiser Permanente, Oakland, CA 94612, USA.

*To whom correspondence should be addressed. E-mail: kathleen.merikangas@nih.gov

tering of multiple sclerosis is almost totally attributable to shared genes (*6*). Although the λ value for breast cancer is of relatively low magnitude, empirical evidence implicates a genetic basis for the familial aggregation and age of onset of most cancers (*7*).

For some disorders, such as nicotine dependence, there is a dearth of empirical data on multigenerational patterns of transmission. On the basis of heritability estimates from twin studies alone, numerous investigators are searching for genes for smoking. The results of recent family and twin studies have revealed that the spouse correlations for nicotine dependence were equivalent to sibling correlations, a finding inconsistent with virtually any genetic model (*8*, *9*). Moreover, the finding that peer influence far outweighs sibling smoking (*9*, *10*) and the low parent-offspring concordance on adolescent smoking (*11*, *12*) further suggest that research should focus on the social transmission of smoking. Large-scale gene searches that fail to integrate these potent environmental contributions are likely to come up empty-handed.

The second column of Table 1 presents susceptibility genes that have been reproducibly identified by independent investigations. Most of these genes explain a small proportion of cases in the general population. For example, the BRCA1 and BRCA2 loci account for about 5 to 10% of breast cancer cases in the population (*13*). The role of susceptibility genes for Alzheimer's disease illustrates the importance of the concept of population-attributable risk, the proportion of disease incidence that is attributable to a specific risk factor, in determining the public health significance of a particular risk factor (*14*). Three major genes have been identified (β-amyloid precursor protein, presenilin-1, and presenilin-2) that are deterministic (i.e., nearly completely penetrant) for the development of Alzheimer's disease in the very rare families who harbor these mutations (*15*). The far more common apolipoprotein E (apoE) ε4 allele increases the odds of Alzheimer's disease by 3 times among those with one copy and by 15 times among those with two copies of this gene (*16*). Because of its high population prevalence, the apoE ε4 allele has a far greater population-attributable risk than the less frequent but extremely high-risk genes described above.

Have Specific or Nonspecific Environmental Factors Been Identified?

As shown in Table 1, there is abundant evidence for the significance of environmental contributions to complex diseases. A specific environmental exposure is prerequisite to several of the disorders shown in the table, including acquired immunodeficiency syndrome (AIDS), alcohol and nicotine dependence, and cervical cancer (*17*, *18*). The marked increase in the incidence of several complex disorders during the past few decades and pronounced regional differences in population prevalence (e.g., type 2 diabetes, AIDS, and alcohol and nicotine dependence) strongly implicate environmental factors (*19*, *20*). For example, there is a notable (i.e., fivefold) elevation in obesity and type 2 diabetes in U.S. Pima Indians compared with their Mexican counterparts (*21*). Because the genetic background of these two groups is unlikely to differ (*22*), the differential rates of diabetes must be attributed to environmental variation. Evidence for environmental factors can also be gleaned from migration studies, a particularly powerful tool to identify environmental exposures that may modify gene expression. Migration studies have shown that the rates of breast cancer, heart disease, and type 2 diabetes have increased substantially in several subgroups as they migrated to new cultural settings (*23*–*25*).

Intervention studies are one of the most potent methods to test environmental etiology. Systematic trials of weight loss and exercise among high-risk offspring of parents with type 2 diabetes have shown a significant decrease in incidence among those in the weight loss and exercise groups (*26*). The increased incidence of type 2 diabetes among offspring exposed prenatally to maternal diabetes also suggests that intrauterine exposure should be incorporated into genetic studies of this disorder (*27*).

The malleability of environmental risk factors is a particularly important consideration in determining priorities for public health. Fixed factors, such as sex, birth cohort, and ethnicity, may be important in characterizing risk but cannot serve as targets of prevention. The major preventable environmental causes of death or illness are: tobacco use, unhealthy diet, physical inactivity, excessive alcohol use, infections, trauma, and exposure to environmental toxins (*28*, *29*).

The dramatic reduction in smoking in the general population of the United States has been one of the most important recent public health successes. In a relatively short time, the Marlboro man has been relegated to glass booths in airports. It is likely that widespread social changes including increased cigarette prices, media

Table 1. Characteristics of Selected Complex Diseases

Disease	Phenotype measure	Genes		Specific environmental factors		Prevalence	Impact (DALY)	Ref.
		λ	Confirmed loci	Known?	Malleable			
Breast cancer	Biopsy	1.8	BRCA-1 BRCA-2	Parity, 1st child > age 30, physical inactivity	Possible	1.2%	0.4	(*13*, *18*)
Alzheimer's disease	Clinical, neurocognitive testing, postmortem biopsy	2.8	PS1/PS2 APP APOE	Head injury, low educational level	Possible	5% (>age 65)	0.8	(*14*, *15*, *41*)
Type 1 diabetes	Immunologic markers, glucose metabolism	15.0	HLA INS	Nonspecific	No	0.4%	0.1	(*17*, *42*, *43*)
Multiple sclerosis	Clinical, neuroimaging	20.0	HLA	Nonspecific	No	0.2%	0.1	(*6*)
Autism	Clinical	60.0	None	No	No	.02%	No info	(*44*)
Schizophrenia	Clinical	9.0	None	No	No	0.8%	1.1	(*45*)
Cervical cancer	Biopsy	1.8	None	Human papilloma virus	Yes	0.16%	0.3	(*18*)
Type 2 diabetes	Glucose metabolism	4.3	PPARγ	Obesity, physical inactivity	Yes	6.1% (>age 20) 15.0% (>age 60)	1.1	(*19*, *46*)
AIDS	Clinical, antibody, $CD4^+$ count	NA	CCR5 HLA	HIV	Yes	0.12%	6.0	(*47*, *48*)
Nicotine dependence	Interview	1.4	None	Nicotine	Yes	24.0%	–	(*9*, *49*)
Alcohol dependence	Interview	7.0	ADH2 ALDH2 (protective)	Alcohol	Yes	4.0%	4.2	(*50*, *51*)

campaigns, and community interventions played a far more important role in inducing this change than programs that target specific risk factors for smoking (*30*). The potent effect of pervasive societal changes on this behavior will far outweigh any possible benefits of identification of risk genes for individual smokers (*31*, *32*).

What Is the Public Health Impact?

The public health impact of a disease depends on the population prevalence and its impact on disability or impairment in occupational and social functioning (Table 1). By contrast, interventions focused on genes associated with very rare diseases will have little impact on public health, whereas even a small reduction in risk for common diseases may have major public health significance. To compare the relative disability of each of the exemplary diseases, we present the disability-adjusted life years (DALYs), defined as the percentage of total years of life lost to disability or premature death for the diseases in Table 1 (*33*). The concept of attributable risk may also be applied to estimate the percent of DALYs that could be reduced by preventing exposure to particular genetic or environmental risk factors. The diseases in Table 1 with the greatest DALYs, AIDS and alcohol dependence, are those that are most amenable to environmental modification. Research designed to decrease exposure or increase protection from these environmental exposures would have dramatic impact on population DALYs for chronic diseases.

Conclusions

The ultimate goal of all medical research is the reduction of morbidity and mortality through prevention and treatment. The translation of genomics to human disease will most likely involve genetic counseling, drug therapy, and gene therapy, once the technology for efficacious application of the latter approaches has been established. These interventions will be most effective for rare diseases at the level of the individual or specific families. In contrast, public health prevention is most effective when applied to common diseases at the population level (*34*). The prevention of phenylketonuria through newborn screening is one example of the successful integration of genetics and public health (*35*). Meanwhile, there is a growing consensus that widespread genomic profiling will not be useful until the full potential of genomics and its intersection with public health has been more fully exploited (*36*–*39*). When making choices in research, the potential for advancing our understanding of human biology, insight into disease pathogenesis, and the impact of those findings on the individuals and families affected by rare, severe diseases must always be balanced with the public health impact of ameliorating more common diseases with lower mortality but high levels of disability.

We propose that the expensive and laborious tools of molecular genetics be prioritized to those diseases that have implicated genes for the familial aggregation and that cannot now be treated or prevented with environmental changes. Examples in Table 1 include breast cancer, Alzheimer's disease, type 1 diabetes, multiple sclerosis, autism, and schizophrenia. In contrast, gene hunting for disorders that appear to be highly amenable to environmental modification, such as type 2 diabetes, AIDS, alcohol dependence, and nicotine dependence, would have lower priority, even though genes may be involved in their etiology. Resources would be far better placed in designing effective interventions and studying the causes of the gap between knowledge and modification of health-related behaviors.

Even more important in considering priorities is the role of personal choice of health behavior in elevating or decreasing disease risk (*40*). Sedentary life styles and dietary choices are important risk factors for type 2 diabetes and cardiovascular diseases, and drug use and sexual behavior are critical in the transmission of the human immunodeficiency virus (HIV). Consumption of alcohol and cigarettes is prerequisite to the development of the addictions. Efforts to prevent use or exposure would be far more effective than treatment of these severe conditions once they progress, even if gene-based treatment strategies should emerge.

We need to study behavioral motivation and effective methods for inducing changes in human behavior before we can even begin to envision genetic counseling for disorders such as nicotine or alcohol dependence. It is also unlikely that knowledge of the genes that may accelerate their consequences will alter this situation. Although these behaviors may have some genetic underpinnings, we are still a long way from identifying them.

In conclusion, identification of genes through the powerful technology of genomics is likely to have a greater impact on Mendelian diseases and some complex diseases than on others. We propose several criteria that may be used to establish priorities for genomics research. Complex diseases with the strongest genetic contribution, limited ability to modify exposure or risk factors, and high public health impact should have the highest priority for genetics research. For those disorders with which genomics tools may have less impact, public health approaches may ultimately lead to far more effective prevention and intervention.

References and Notes

1. A. M. Glazier, J. H. Nadeau, T. J. Aitman, *Science* **298**, 2345 (2002).
2. J. Little *et al.*, *Am. J. Epidemiol.* **157**, 667 (2003).
3. P. A. Baird, *Perspect. Biol. Med.* **45**, 516 (2002).
4. I. Gottesman, J. Shields, *Schizophrenia and Genetics: A Twin Study Vantage Point* (Academic Press, New York, 1972).
5. N. Risch, *Am. J. Hum. Genet.* **46**, 222 (1990).
6. A. D. Sadovnick, *Clin. Neurol. Neurosurg.* **104**, 199 (2002).
7. N. Risch, *Cancer Epidemiol. Biomarkers Prev.* **10**, 733 (2001).
8. D. I. Boomsma, J. R. Koopmans, L. J. P. Van Doornen, J. F. Orlebeke, *Addiction* **89**, 219 (1994).
9. K. R. Merikangas, L. Dierker, S. Avenevoli, H. Zhao, N. Risch, *Am. J. Epidemiol.*, in preparation.
10. S. Avenevoli, K. R. Merikangas, *Addiction* **98** (suppl. 1), 1 (2003).
11. J. S. Rose, L. Chassin, C. C. Presson, S. J. Sherman, in *Merrill Palmer Quarterly* (Wayne State Univ. Press, Detroit, MI, 1999), vol. 45, pp. 62–84.
12. K. Kobus, *Addiction* **98** (suppl. 1), 37 (2003).
13. M. L. Slattery, R. A. Kerber, *JAMA* **270**, 1563 (1993).
14. A. Slooter, C. van Duijn, *Epidemiol. Rev.* **19**, 107 (1997).
15. R. L. Nussbaum, C. E. Ellis, *N. Engl. J. Med.* **348**, 1356 (2003).
16. L. A. Farrer *et al.*, *JAMA* **278**, 1349 (1997).
17. H. K. Akerblom, O. Vaarala, H. Hyoty, J. Ilonen, M. Knip, *Am. J. Med. Genet.* **115**, 18 (2002).
18. American Cancer Society, "Cancer facts and figures 2003" (American Cancer Society, Atlanta, GA, 2003).
19. S. J. Kenny, R. E. Aubert, L. S. Geiss, in *Diabetes in America* (National Diabetes Data Group, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, NIH publication no. 95-1468, Bethesda, MD, ed. 2, 1995), pp. 47–68.
20. M. R. Rewers, R. F. Hamman, in *Diabetes in America* (National Diabetes Data Group, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, NIH publication no. 95-1468, Bethesda, MD, ed. 2, 1995), pp. 179–220.
21. E. Ravussin, M. E. Valencia, J. Esparza, P. H. Bennett, L. O. Schulz, *Diabetes Care* **17**, 1067 (1994).
22. K. Kidd, A. J. Pakstis, W. C. Speed, J. R. Kidd, personal communication.
23. R. Trevisan, M. Vedovato, A. Tiengo, *Nephrol. Dial. Transplant.* **13** (suppl. 8), 2 (1998).
24. A. J. Hanley, B. C. Choi, E. J. Holowaty, *Int. J. Epidemiol.* **24**, 255 (1995).
25. A. G. Shaper, J. Elford, *Acta Paediatr. Scand. Suppl.* **373**, 73 (1991).
26. W. C. Knoller *et al.*, *N. Eng. J. Med.* **346**, 393 (2002).
27. E. Sobngwi *et al.*, *Lancet* **361**, 1861 (2003).
28. E. J. Calabrese, E. J. Stanek, R. C. James, S. M. Roberts, *Environ. Health Perspect.* **105**, 1354 (1997).
29. N. Ishibe, K. T. Kelsey, *Cancer Causes Control* **8**, 504 (1997).
30. P. M. Lantz *et al.*, *Tob. Control* **9**, 47 (2000).
31. G. Rose, *The Strategy of Preventive Medicine* (Oxford Univ. Press, London, 1997).
32. M. Susser, E. Susser, *Am. J. Public Health* **86**, 668 (1996).
33. World Health Organization, "World health report 2002: Reducing risk, promoting healthy life" (World Health Organization, Geneva, 2002).
34. M. Susser, *J. Epidemiol. Community Health* **52**, 608 (1998).
35. R. J. Pollitt, *J. Inherit. Metab. Dis.* **24**, 299 (2001).
36. S. B. Haga, M. J. Khoury, W. Burke, *Nature Genet.* **34**, 347 (2003).
37. M. J. Khoury, the Genetics Working Group, *Am. J. Public Health* **86**, 1717 (1996).
38. M. J. Khoury, *Genet. Med.* **5**, 261 (2003).
39. P. Baird, *Community Genet.* **4**, 77 (2001).
40. S. Shostak, *Soc. Sci. Med.* **56**, 2327 (2003).
41. H. C. Hendrie, *Am. J. Geriatr. Psychiatry* **6**, 3 (1998).
42. F. Pociot, M. F. McDermott, *Genes Immunol.* **3**, 235 (2002).
43. S. S. Rich, P. Concannon, *Diabetes* **51** (suppl. 3), S288 (2002).
44. C. J. Newschaffer, D. Fallin, N. L. Lee, *Epidemiol. Rev.* **24**, 137 (2002).
45. E. M. Goldner, L. Hsu, P. Waraich, J. M. Somers, *Can. J. Psychiatry* **47**, 833 (2002).
46. T. A. Harrison *et al.*, *Am. J. Prev. Med.* **24**, 152 (2003).
47. Centers for Disease Control, "Centers for Disease Control and Prevention, National Center for HIV, STD, and TB Prevention, Division of HIV/AIDS Prevention—Surveillance and Epidemiology, AIDS Surveillance, special data run" (Centers for Disease Control, Atlanta, GA, 2001).
48. M. Carrington, S. J. O'Brien, *Annu. Rev. Med.* **54**, 535 (2003).
49. N. Breslau, E. O. Johnson, E. Hiripi, R. Kessler, *Arch. Gen. Psychiatry* **58**, 810 (2001).
50. K. R. Merikangas, *Psychol. Med.* **20**, 11 (1990).
51. D. Goldman, in *Recent Developments in Alcoholism: Ten Years of Progress*, M. Galanter, Ed. (Plenum Press, New York, 1993), vol. 2.