The crucial role of the public health sciences in the postgenomic era

Gilbert S. Omenn, MD, PhD

Genomic sequence information and gene and protein expression patterns must be linked with information about diet and metabolism, lifestyle behaviors, diseases and medications, and microbial, chemical, and physical exposures. These linkages depend upon the public health sciences: epidemiology, biostatistics, environmental health sciences, pathobiology, health sciences research, and clinical prevention trials. *Genet Med* 2002:4(6, Supplement):21S–26S.

Key Words: *genomics, public health sciences, risk assessment, toxicogenomics, host–parasite interactions*

We are changing the world in which we live through science. The New Biology and its associated technologies of gene expression microarrays, comparative genomics, proteomics, and bioinformatics are the core scientific material of this Symposium. In 2000, *Science* magazine declared the leading breakthrough in all of science to be "the sequenced genomes."[1] These genomes included not just the human, but also yeast, *Escherichia coli*, and other microorganisms, the fruit fly *Drosophila melanogaster*, the worm *C. elegans*, and the plant *Arabidopsis*. Together, the homologies among these sequences demonstrate the unity of biological processes across diverse species. This principle that crucial biological information is shared across species gives us great power for identifying and confirming the biochemical functions of genes and their action that might be important in health and disease. This work is not the end, not even the beginning of the end, but, perhaps, the end of the beginning.

As Francis Collins, Director of the United States National Human Genome Research Institute and a faculty member on leave from the University of Michigan, has written: "Mapping the human genetic terrain may rank with the great expeditions of Lewis and Clark, Sir Edmund Hillary, and the Apollo Program."[2]

Public and media interest was intense when President Clinton and Prime Minister Tony Blair jointly announced the broad results of accelerated progress by the public and private sector sequencing programs in June 2000. The details of those "blueprints" for the human genome sequence were published, along with many interpretive articles, in *Nature* and in *Science* in mid-February 2001.[3,4] Considerable additional work is ongoing to resolve uncertainties in the placement of many fragments, especially those involving duplicated sequences, and in

deciphering the sequences in the centromeric regions of the chromosomes. One stunning finding was the much lower number of "genes" identified and deduced, perhaps 30,000 to 40,000, instead of the long-estimated 50,000 to 100,000. Nevertheless, splicing and alternative transcription generate at least the expected approximately 100,000 protein products from this more constrained number of genes.

We are very far from fully understanding all the biological information in the human genome. The functional challenge before us is to link massive quantities of sequence information from the genome databases to the function of each gene product in health and disease. It is far too simplistic, especially for common diseases, to associate individual genes and individual variants or single nucleotide polymorphisms of genes with disease risk; it is essential to know the other genes and the many nongenetic factors that contribute to increasing or decreasing the risk of each disease in individuals and in population groups.

A GOLDEN AGE FOR THE PUBLIC HEALTH SCIENCES

Genomic sequence information must be linked with information about nutrition and metabolism, lifestyle behaviors, diseases and medications, and microbial, chemical, and physical exposures.[5–7] I will emphasize in this article the environmental–genetic interactions. Table 1 highlights the roles of each of the public health sciences in the postgenomic era.

We may note that genetics and public health share certain salient attributes. Both focus on populations. Both need more information about heterogeneity—of genetic predispositions, of environmental exposures, of disease risks, and of responses to interventions—within and across population subgroups. Both explicitly recognize the importance of cultural, societal, ethnic, and racial contexts. And both are sensitive about the legacy and the risks of discrimination on social and racial grounds.

In the clinical setting, genetics provides a bridge between medicine and public health. Counseling and treatment of individual patients must often be expanded to nuclear or ex-

Table 1
The postgenomic era will depend upon the public health sciences

- Epidemiology

    All the factors that play upon/determine the expression of genetic variation identified in sequences; useful roles of biomarkers

- Biostatistics and bioinformatics

    Platforms and methods for designing studies and for analyzing the avalanche of data

- Environmental health sciences/"ecogenetics"

    Understanding the host variation in host–agent interactions for risk assessment/risk management

- Pathobiology/infectious disease

    Host–pathogen genomic interactions

    Host–pathogen environmental interactions

- Behavioral sciences

    Genetic predispositions to aspects of smoking behavior and other unhealthful behaviors

- Health services research fields

    Designing and assessing well-targeted, cost-effective genetic, clinical, and preventive services that improve quality of life

tended families. Monitoring and screening for genetic predispositions or genetically predisposed diseases involves worker populations and sometimes general population groups, while the stimulus for such monitoring and screening may come from care of individual patients. Outreach to the likely affected communities should involve not just earlier diagnosis, but also, very importantly, prevention. Both for individual patients and for population groups, specific biochemical or chromosomal tests can change probabilistic genetic counseling to diagnostically specific advice for particular conditions in particular individuals. Finally, the accelerating pace of discoveries and applications accentuates the need for education about genetics for health professionals and about public health for geneticists.[8,9]

Epidemiology

Biochemical and molecular epidemiology is a bridge between public health serving populations and clinical medicine serving individuals. During the past decade, there has been a remarkable transformation of epidemiology. At the front end, there is a dramatic move beyond statistical associations to test hypotheses of mechanisms of disease (etiological research). Use of biomarkers of exposure, of effect, and of susceptibility not only refers to potential mechanisms but also provides measurable links between experimental observations in animal models and clinical research observations in people.[10] There is a huge need to couple laboratory and epidemiological approaches. The power of these studies is being greatly expanded with the introduction of gene expression and protein expression methods, as noted in this Symposium volume. With re-

gard to cancers and other diseases with long latency until symptoms appear, the use of biomarkers and intermediate clinical or pathological endpoints can reduce the latent period and therefore the duration of study required to draw inferences and propose causal relationships.

Once etiological hypotheses have been generated and tied to credible potential mechanisms, investigators can attempt to validate the hypotheses by modifying or removing risk factors in prevention clinical trials. Some trials involve behavior change, such as efforts to increase the rates of smoking cessation. Others use pharmaceuticals, vitamins, or natural products for chemoprevention of certain diseases. Examples are antioxidant vitamins (beta-carotene, C, E), vitamin A and other retinoids, folic acid, and inducers of glutathione.[11,12] For infectious agents, vaccines can be highly effective preventive interventions. And for environmental and occupational risk reduction, it is feasible to monitor emissions, exposures, and subclinical and clinical effect rates after actions are taken.

Biostatistics and bioinformatics

Platforms and specific software for a variety of study designs and data analyses are essential for modern genetic and genomic research. Just as much major laboratory work has been automated and robotized, much statistical work of creating databases, acquiring and cleaning up data, and analyzing findings has been automated with computerized programs.[13,14] The interaction of mathematics and biology, computational biology, holds great promise.

Environmental health sciences

The realization has finally taken hold that debates that pit "genetic" versus "environmental" or "nature" versus "nurture" as causes of various diseases are inappropriate. The action lies in the interaction of these polar views, leading to the emergence during the past 30 years of the subfield of "ecogenetics."[15,16]

Ecogenetics

Besides the scientific logic of investigating interactions between genetic variation in exposed people or other species (the host) and variation in exposure to specific infectious or chemical agents in particular environments, there is a strong regulatory rationale for ecogenetics studies. The United States Occupational Safety and Health Act of 1972 mandated that health standards be set "such that no worker shall suffer adverse effect. . ." if exposed at the maximal permissible level for a working lifetime.[17] Physicians seeing patients with workplace-related clinical conditions experience a more direct challenge to understand host variability in susceptibility. Often a patient, told that the symptoms may be due to exposures on the job, will ask, "Why me, Doc? I'm no less careful than the next guy." The Clean Air Act Amendments of 1977 required that section 109 criteria air pollutant standards be set "so as to protect the most susceptible subgroup in the population."[17] That proviso can be met only if there are studies to define the most suscep-

tible subgroup and the levels of exposure that are hazardous for that subgroup. An example might be persons with cystic fibrosis at risk for lung impairment from chronic exposures to elevated ozone (photochemical oxidant) in inhaled air. Instead, the 1979 revision of the ozone standard was based on the susceptibility of the large population subgroup with asthma, bronchitis, or emphysema (3–5% of the general population). Finally, the Food Quality Protection Act of 1996 requires federal regulators at the EPA and FDA to address risks for vulnerable or unusually exposed subgroups. In response, the EPA has put special attention on estimated risks of pesticide exposures for children.[18]

Risk assessment

Already we have used the term "risk" multiple times, referring to the probability of adverse effects in particular groups of people experiencing potentially hazardous exposures. In 1997, a Presidential/Congressional Commission on Risk Assessment and Risk Management issued its final reports.[19] Its six-stage Framework for Risk Management, shown in Figure 1, placed risk assessment in the broader context of risk management. The Framework had two especially noteworthy features. First, the Commission made engagement of affected stakeholders central and urged that such engagement be initiated promptly at the start of the process rather than waiting until frustrations build and mistrust is accentuated by perceived exclusion during the lengthy technical phases. The Commission presented numerous examples in which community-based stakeholders raised issues for assessment that might otherwise have been neglected or proposed practical solutions for risk reduction that might otherwise have been rejected by experts and policy makers not willing to put any burden on the exposed popula-

tions to modify exposures.[19] Second, the Commission urged that each new or newly salient environmental problem be put into a broader public health context, giving the public and the agency experts a much better sense of the nature of the adverse effects on health or ecosystems and the relative risk compared with other known risks from agents with similar effects. Putting problems into context requires examination of multiple sources of the same chemical, pathways of exposure to the chemical through multiple media, other causes of the same endpoint(s) (generating an estimate of the attributable fraction for the agent under review), and multiple effects of the same chemical.

The scientific disciplines, analyses, and potential actions for hazard identification, risk assessment, and risk management are outlined in Figure 2, which comes from a 1980 publication of the White House Office of Science and Technology Policy,[20] reinforced by the 1983 report from the National Research Council, called "Risk Assessment in the Federal Government: Managing the Process."[21]

Toxicogenomics

Led by the National Institute for Environmental Health Sciences of the United States National Institutes of Health, the federal government has mobilized to develop molecular signatures for exposures, early effects, and variation in susceptibility to chemical agents that cause cancers, mutations, birth defects, and organ system dysfunction. One way to jump start this field is to test known carcinogens for distinctive patterns of gene and protein expression by using such agents as benzidines, β-naphthylamine, benzene, bis-chloromethylether, nitrosamines, and asbestos in animal models. Similarly, it is quite feasible to seek molecular signatures of cancer chemopreventive agents in animal models and in organ and cell cultures.

A ToxChip has been developed with a custom human cDNA microarray starting from 750,000 sequences at GenBank, and 65,000 nonredundant clusters in UniGene, ending with 2,090 unique human genes. Among these genes, 72 are known or postulated to function in apoptosis, 90 in oxidative stress–redox homeostasis, 22 in peroxisome proliferation responses, 12 in the Ah receptor battery, 84 in housekeeping functions, 63 estrogen responsive, 76 oncogenes and tumor suppressor genes, 51 in cell cycle control, 131 as transcriptional factors, 276 kinases, 88 phosphatases, 23 heat shock proteins, 30 cyto-
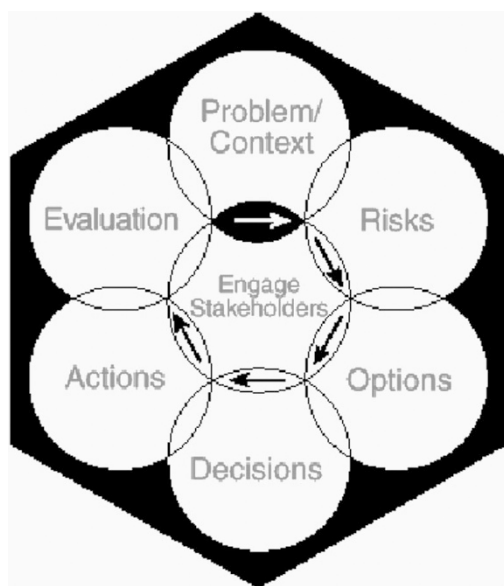
Fig. 1 Risk Commission ("Omenn Commission") six-stage Framework for Risk Management.[19] The critical role of stakeholders in setting the context and guiding technical assessments is indicated by the larger ellipse in Stage 1. The arrow is removed from Stage 6 so as not to encourage "paralysis by analysis."
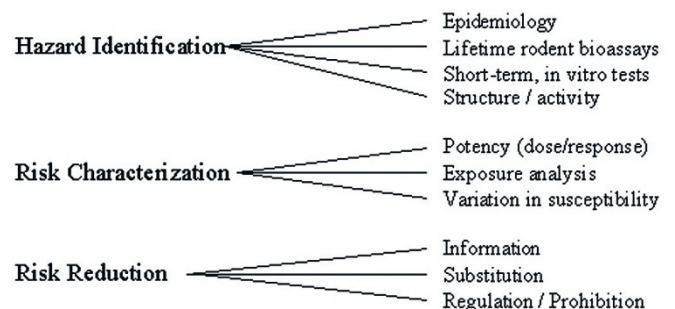
Fig. 2 Framework for Regulatory Decision-Making.[20]

chrome P450s, and 349 as receptors.[22] During 1999 to 2002, the National Institute of Environmental Health Sciences Environmental Genome Project resequenced 123 of its list of 554 environmentally responsive genes and identified more than 1,700 SNPs in these genes (work at University of Washington and University of Utah Genome Centers, www.niehs.nih.gov/envgenom).[22] It is likely that some of these methods will become common in the next few years in reference laboratories for pharmacology and toxicology.

Pathobiology: Infectious diseases and host–parasite interactions

There are prominent examples of ecogenetic relationships between variation in susceptibility and agents of malaria (*Plasmodium falciparum*, *Plasmodium vivax*), tuberculosis (*Mycobacterium tuberculosis*), AIDS (human immunodeficiency virus), cholera (*Vibrio cholerae*), and meningitis–otitis (*Haemophilus influenzae*).[23,24] As of spring 2001, prokaryotic genome sequences were complete for 39 bacteria and 9 Archaebacteria; at least 100 more species were substantially sequenced, with annotations in progress.[25]

Genomics has accelerated insights about microorganisms, including genome architectural features (genic content and repetitive sequences), sequence similarities (orthologs and paralogs, protein and DNA motifs), mobile genetic elements (phage, pathogenicity islands), and large numbers of genes of previously unknown function. An illustrative organism is *H. influenzae*. This organism was the first to be sequenced, a stunning achievement published in 1995.[26] Its advantages for research include its small size of 1830 kb, its importance as a human pathogen with a mouse model, its capacity for DNA transformation, and rapidly advancing knowledge of its genome. There are 1703 proposed genes, of which 736 lack proposed functions; of these, 347 are conserved across species, while 389 are unique to *H. influenzae*.[27] The unique genes are preliminary targets for new therapeutic agents and new vaccines. Targets that are essential genes cannot be lost by deletion to generate resistance and are unlikely to undergo phase variation or antigenic variation, which have undermined vaccines for certain viral diseases (influenza, HIV). Akerley has developed a method called genomic analysis and mapping by in vitro transposition [genomic analysis and mapping by in vitro transposition (GAMBIT)] to identify essential and conditionally essential genes in transformable microorganisms. GAMBIT combines long-range polymerase chain reaction, high efficiency in vitro transposition, transformation and recombination, and genetic footprinting. A key next challenge is performing a complete proteome analysis of a major human pathogen, using yeast 2-hybrid screening for interactions, proteome maps of tagged protein complexes, and complementary computational methods.

Genomes of many pathogenic microbes and their innocuous relatives have now been sequenced, with important insights. The genome size and gene number vary widely, reflecting diversity in ecological niches. In general, organisms that need to survive diverse environments have larger genomes with comprehensive biosynthetic pathways. In contrast, obligate parasites have smaller genomes with adaptations that facilitate an existence entirely dependent on the host. For example, *M. tuberculosis* has genomic expansions of enzymes involved in lipid metabolism and cell wall biogenesis, which facilitate resistance to anti-TB drugs through changes in permeability and transport.[28] This organism uses enzymes of the glyoxylate pathway that seem to enable survival in lung tissue of humans. Meanwhile, *Mycobacterium leprae* is an intracellular obligate parasite with massive gene decay and greatly restricted metabolism, yet it has several species-specific genes and enzymes not found in the larger *M. tuberculosis* genome.[29] Another example is the enterohemorrhagic *E. coli* strain O157:H7, which has expansion of several pathogenic determinants and 1387 additional genes compared with the innocuous K-12 strain.[30] Pathogenicity islands are prominent in pathogens like *V. cholerae*, *Helicobacter pylori*, and *Yersinia pestis*; these genomic regions of 10 to 200 kbp have distinctive structural features and encode adhesins, invasins, toxins, and protein secretion systems that are determinants of bacterial virulence. In the human host, polymorphisms in genes that modulate the immune response in macrophages, cytokines, chemokines, and Toll receptors alter susceptibility to various infections, including HIV[31] and *E. coli* O157:H7 as well as efficacy and safety of antimicrobial drugs. Much work needs to be launched to tie together exposures to microorganisms of varying genetic features with variation in the host genes.

Nutrition and genetics

Genomics and proteomics can help bring modern biology to nutrition, an analogy to the efforts launched in toxicology (See Toxicogenomics). Dietary components induce polymorphic carcinogen-activating and detoxifying enzymes (phase I, phase II biotransformations). Genetic factors are very important in common diseases with substantial dietary influences, such as diabetes mellitus and obesity. Metabolic polymorphisms can influence disease risks, as in the case of a polymorphic variant of methylenetetrahydrofolate reductase (MTHFR), which reduces serum folate levels, permitting higher serum total homocysteine levels and increasing cardiovascular disease risks.[32–34] A program of screening for serum total homocysteine greater than 11 μmol/L and then treating with folic acid has been recommended as cost-effective by Nallamothu et al.[34] Low folate levels are also associated with higher risk of colon cancer in women with a family history of common colon cancers.[35]

A fascinating inherited clinical disorder is hemochromatosis, a common condition (1 per 400 whites) that is characterized by excessive absorption of dietary iron and overload and deposition of iron in various organs. This phenomenon leads to damage of the heart, liver, pancreas, testes, and skin. There is a set of tests for genotype (HFE mutants) and for phenotype (serum transferrin saturation and ferritin levels) and a very simple, inexpensive treatment: periodic venesection. Unfortunately, the large majority of genetically predisposed individuals

Table 2
Principles of community-based research

- Involve community partners from the earliest stages
- Ensure community partners have real influence on project
- Make research processes and outcomes benefit the community
- Invite community members to be part of the analysis/interpretation
- Make productive partnerships last beyond the project
- Empower community members to initiate projects

remain clinically unaffected, so consensus criteria for screening and treatment have yet to emerge.[36,37]

Behavioral sciences

There are clearly genetic predispositions to various aspects of the phenomenon of cigarette smoking and the complications of smoking. The same applies to immediate and long-term organ damage from excessive alcohol intake and probably for other unhealthful behaviors, even physical inactivity.

Health services research

As clinical genetic services are expanded on a population basis, and as many new diagnostic and prognostic genetic tests are introduced, society will need well-organized research on what works and what does not, what is safe and what is not safe, and how best to make useful tests and services cost-effective. In general, we need much more information about the heterogeneity of genetic predispositions and of nongenetic factors and exposures that influence disease risks as well as responses to treatments and preventive interventions in medicine and public health. In pursuing these aims, we must be explicit in anticipating and recognizing cultural, social, ethnic, and racial contexts to avoid discrimination based on genetic and related personal information. Public health research often involves community-based populations. Sensitivity to the interests, aspirations, perceptions, fears, and previous experiences of the community is essential. Table 2 summarizes guidance my colleagues and I have developed in recent years, which may be applied, to varying degrees, universally.

References

1. Pennisi E. Breakthrough of the year: genomics comes of age. *Science* 2000;290:2220–2221.
2. Collins FS. Shattuck lecture: medical and societal consequences of the Human Genome Project. *N Engl J Med* 1999;341:28–37.
3. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860–921.
4. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG et al. The sequence of the human genome. *Science* 2001;291:1304–1351.
5. Omenn GS. Genetics and public health. *Am J Public Health* 1996;86:1701–03.
6. Khoury ML, Burke W, Thompson E. Genetics and public health in the 21st century: using genetic information to improve health and prevent disease. New York: Oxford University Press, 2000.
7. Omenn GS. Public health genetics: an emerging interdisciplinary field for the postgenomic era. *Annu Rev Public Health* 2000;21:1–13.
8. Hager M. Implications of genetics for health professional education. New York: Josiah Macy Jr Foundation, 1999.
9. Austin M, Peyser P, Khoury M. The interface of genetics and public health: research and educational challenges. *Annu Rev Public Health* 2000;21:81–99.
10. National Research Council. Biological markers in pulmonary toxicology. Washington DC: National Academy Press, 1989.
11. Omenn GS. Micronutrients (vitamins, and minerals) as cancer-preventive agents. In: Hakama M, Beral V, Buiatti E, Faivre J, Parkin DM, editors. Principles of chemoprevention. IARC Scientific Publications No. 130. Lyon, France: International Agency for Research on Cancer, 1996:33–45.
12. Omenn GS. Chemoprevention of lung cancer: the rise and demise of beta-carotene. *Annu Rev Public Health* 1998;19:73–99.
13. Lander ES, Stork NJ. Genetic dissection of complex traits. *Science* 1994;265:2037–2048.
14. Rouchka EC, States DJ. Sequence assembly validation by multiple restriction digest fragment coverage analysis. *Proc Int Conf Intell Syst Mol Biol* 1998;6:140–147.
15. Brewer GJ. Annotation: human ecology, an expanding role for the human geneticist. *Am J Hum Genet* 1971;23:92–94.
16. Omenn GS, Motulsky AG. Ecogenetics: genetic variation in susceptibility to environmental agents. In: Cohen BH, Lilienfeld AM, Huang PC, editors. Genetics issues in public health and medicine. Springfield, IL: CC Thomas, 1978:83–111.
17. Omenn GS, Faustman EM. Risk assessment and risk management. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. New York: Oxford University Press, 2002:1083–1103.
18. Goldman LR. Linking research and policy to ensure children's environmental health. *Environ Health Perspect* 1998;106:857–862.
19. Presidential/Congressional Commission on Risk Assessment, and Risk Management. Volumes I, and II, Final Report. Washington DC: Government Printing Office, 1997 [Available on-line at www.riskworld.com, under Reports, confirmed 5/28/02].
20. Calkins DR, Dixon RL, Gerber CR, Zarin D, Omenn GS. Identification, characterization and control of potential human carcinogens: a framework for Federal decision making. *J Natl Cancer Inst* 1980;64:169–175.
21. National Research Council. Risk assessment in the federal government: managing the process. Washington DC: National Academy Press, 1983.
22. The National Institute of Environmental Health Sciences, 2002 [www.niehs.nih.gov/envgenom]. The genomic era: a crucial role for the public health sciences. *Environ Health Perspect* 2000;108:160–161.
23. McNicholl J, Downer MV, Udhayakumar V, Alper CA, Swerdlow DL. Host-pathogen interactions in emerging and re-emerging infectious diseases: a genomic perspective of TB, malaria, HIV, hepatitis B, and cholera. *Annu Rev Public Health* 2000;21:15–46.
24. Miller LH. Impact of malaria on genetic polymorphism and genetic diseases in Africans and African-Americans. *Proc Natl Acad Sci U S A* 1997;91:2415–2419.
25. Subramanian G, Adams MD, Venter JC, Broder S. Implications of the human genome for understanding human biology and medicine. *JAMA* 2001;286:2296–2307.
26. Fleischman RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb J-F, Dougherty BA, Merrick JM, McKenney K, Sutton G, FitzHugh W, Fields C, Gocayne JD, Scott J, Shirley R, Liu L, Glodek A, Kelley JM, Weidman JF, Phillips CA, Spriggs T, Hedblom E, Cotton MD, Utterback TR, Hanna MC, Nguyen DT, Saudek DM, Brandon RC, Fine LD, Fritchman JL, Fuhrmann JL, Geoghagen NSM, Gnehm CL, McDonald LA, Small KV, Fraser CM, Smith HO, Venter JC. Whole-genome random sequencing and assembly of haemophilus influenzae Rd. *Science* 1995;269:496–512.
27. Akerley BJ, Rubin EJ, Novick VL, Amaya K, Judson N, Mekalanos JJ. A genome-scale analysis for identification of genes required for growth or survival of Haemophilus influenzae. *Proc Natl Acad Sci U S A* 2002;99:966–971.
28. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE III, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, OsborneJ, Quail MA, Rajandream M-A, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998;393:537–544.
29. Cole ST, Eiglmeier K, Parkhill J, James K, Thomson NR, Wheeler PR, Honoré N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chilllingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream M-A, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG. Massive gene decay in the leprosy bacillus. *Nature* 2001;409:1007–1011.
30. Perna NT, Plunkett G III, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Pósfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamousis KD, Apodaca J, Ananthara-

man TS, Lin J, Yen G, Schwartz DC, Welch RA, Blattner FR. Genome sequence of enterohaemorrhagic *Escherichia coli* 0157:H7. *Nature* 2001;409:529–533.

31. Littman DR. Chemokine receptors: keys to AIDS pathogenesis? *Cell* 1998;93:677–680.

32. Boushey C, Beresford SAA, Omenn GS, Motulsky AG. A quantitative assessment of plasma homocysteine as a risk factor for vascular disease: probable benefits of increasing folic acid intakes. *JAMA* 1995;274:1049–1057.

33. Omenn GS, Malinow MR, Barnett M, Thornquist MD, Goodman GE, and the CARET Study Group. A prospective study of hyperhomocystinemia and low folic acid as risk factors for cardiovascular mortality (unpublished).

34. Nallamothu BK, Fendrick AM, Rubenfire M. Saint S, Bandekar R, Omenn GS. Potential clinical and economic effects of homocysteine lowering. *Arch Int Med* 2000;160:3406–3412.

35. Fuchs CS, Willett WC, Colditz GA, Hunter DJ, Stampfer MJ, Speizer FE, Giovannucci EL. The influence of folate and multivitamin use on the familial risk of colon cancer in women. *Cancer Epidemiol Biomarkers Prev* 2002;11:227–234.

36. Motulsky AG, Beutler E. Population screening in hereditary hemochromatosis. *Annu Rev Public Health* 2000;21:65–79.

37. Cogswell ME, Burke W, McDonnell SM, Franks AL. Screening for hemochromatosis: a public health perspective. *Am J Prev Med* 1999;16:134–140.