

antibodies or oligonucleotide aptamers. Their potential applications include: revealing interactions among proteins and between proteins and small molecules (drugs) or other ligands, identifying substrates for a modifying enzyme such as a protein kinase, and searching for enzymatic activities. A harbinger of the promise of this approach is the recent demonstration of proteins in nanoliter droplets immobilized by covalent attachment to glass slides; more than 10,000 samples could be spotted onto each slide with this technique (31). The few test proteins in this array format were assayed for interactions with another protein or a small molecule, and for their phosphorylation by a protein kinase. Targeted arrays will allow the identification of all of the enzymes in an organism that are able to carry out a specific modification of a substrate; for example, protein arrays have tested nearly the entire set of the predicted protein kinases in yeast for their activity on 17 substrates (32).

Protein databases will need to become much more sophisticated if they are to help scientists make sense of the staggering number of experimental measurements that will soon emerge. Demands range from tracking all of the ligands for each analyzed member of a protein family (such as the SH3 domain) to cataloging all of the known substrates of each protein kinase, protein phosphatase, or other modifying activity. In addition, protein data will need to be integrated with results from expression profiling, genome-wide mutation or antisense analyses, and polymorphism detection. As proteomic data accumulate, we will become better at triangulating from multiple disparate bits of information to gain a bearing on what a protein does in the cell. Proteomics will come of age when its revelations about formerly uncharacterized proteins directly drive imaginative hypotheses about their functions.

What We Need

For a field so laden with razzmatazz methods, it is striking that the number one need in proteomics may be new technology. There are simply not enough assays that are sufficiently streamlined to allow the automation necessary to perform them on a genome's worth of proteins. Those currently available barely scratch the surface of the thousands of specialized analyses biologists use every day on their favorite proteins. What we need are experimental strategies that could be termed cell biological genomics, biophysical genomics, physiological genomics, and so on, to provide clues to function. In addition, a protein contains so many types of information that each of its properties needs to be assayed on a proteome-wide scale, ideally in a quantitative manner.

As we have argued (33), existing technology—and more importantly, the reagents

(sets of genes, plasmids, strains, proteins, and the like) and equipment to handle these reagents—must rapidly spread from the specialized genomic and proteomic centers to the rest of the community. Only when every laboratory is comfortable doing proteomics will its power be exploited fully. Moreover, the likelihood of new approaches increases in proportion to the number of investigators participating in the field.

An interdisciplinary spirit will come to guide those excited by the global analysis of protein function. Geneticists need to talk to chemists, physiologists to physicists, cell biologists to computer scientists. With questions so grand, the expertise to answer them requires the entire spectrum of science. This combination of new technology and its widespread dispersion together with broad-ranging collaborative projects will culminate in the frabjous day when the undertaking that began with genome sequencing reaches fruition.

References and Notes

1. S. P. Gygi, Y. Rochon, B. R. Franza, R. Aebersold, *Mol. Cell. Biol.* **19**, 1720 (1999).
2. W. Blackstock, in *Proteomics: A Trends Guide*, W. Blackstock, M. Mann, Eds. (Elsevier Science, London, 2000), pp. 12–17.
3. G. Neubauer et al., *Nature Genet.* **20**, 46 (1998).
4. M. P. Rout et al., *J. Cell Biol.* **148**, 635 (2000).
5. J. B. Peltier et al., *Plant Cell* **12**, 319 (2000).
6. A. J. Link et al., *Nature Biotechnol.* **17**, 676 (1999).
7. M. P. Washburn, D. Wolters, J. R. I. Yates, *Nature Biotechnol.*, in press.
8. M. Fromont-Racine, J. C. Rain, P. Legrain, *Nature Genet.* **16**, 277 (1997).
9. A. J. Walhout et al., *Science* **287**, 116 (2000).
10. P. L. Bartel, J. A. Roecklein, D. SenGupta, S. Fields, *Nature Genet.* **12**, 72 (1996).
11. S. McCraith, T. Holtzman, B. Moss, S. Fields, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4879 (2000).
12. P. Uetz et al., *Nature* **403**, 623 (2000).
13. B. Schwikowski, P. Uetz, S. Fields, *Nature Biotechnol.*, in press.
14. P. Ross-Macdonald et al., *Nature* **402**, 413 (1999).
15. M. R. Martzen et al., *Science* **286**, 1153 (1999).
16. M. Pellegrini, E. M. Marcotte, M. J. Thompson, D. Eisenberg, T. O. Yeates, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4285 (1999).
17. E. M. Marcotte et al., *Science* **285**, 751 (1999).
18. A. J. Enright, I. Iliopoulos, N. C. Kyriakides, C. A. Ouzounis, *Nature* **402**, 86 (1999).
19. T. Dandekar, B. Snel, M. Huynen, P. Bork, *Trends Biochem. Sci.* **23**, 324 (1998).
20. R. Overbeek, M. Fonstein, M. D'Souza, G. D. Pusch, N. Maltsev, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2896 (1999).
21. R. J. Cho et al., *Molecular Cell* **2**, 65 (1998).
22. P. T. Spellman et al., *Mol. Biol. Cell* **9**, 3273 (1998).
23. S. Chu et al., *Science* **282**, 699 (1998).
24. J. L. DeRisi, V. R. Iyer, P. O. Brown, *Science* **278**, 680 (1997).
25. C. M. Perou et al., *Nature* **406**, 747 (2000).
26. M. Diehn, M. B. Eisen, D. Botstein, P. O. Brown, *Nature Genet.* **25**, 58 (2000).
27. M. L. Bulyk, E. Gentale, D. J. Lockhart, G. M. Church, *Nature Biotechnol.* **17**, 573 (1999).
28. J. G. Hacia et al., *Nature Genet.* **14**, 441 (1996).
29. S. P. Gygi et al., *Nature Biotechnol.* **17**, 994 (1999).
30. O. N. Jensen, in *Proteomics: A Trends Guide*, W. Blackstock, M. Mann, Eds. (Elsevier Science, London, 2000), pp. 36–42.
31. G. MacBeath, S. L. Schreiber, *Science* **289**, 1760 (2000).
32. H. Zhu et al., *Nature Genet.* **26**, 283 (2000).
33. M. Johnston, S. Fields, *Nature Genet.* **24**, 5 (2000).

FUTURE DIRECTIONS: GENOMICS AND MEDICINE

Dissecting Human Disease in the Postgenomic Era

Leena Peltonen and Victor A. McKusick

As overwhelmingly demonstrated by the sequencing papers in this issue, the complete anatomy of the human genome is now before us. In a very short time—within a decade—we have advanced from having very little information about the genetic details of biology to possessing an immense amount of structural information about individual genes. Currently, the complete genome sequences of more than 60 species are available in databases, and the prediction that there will be a total of 100 sequenced genomes in databases within the next few

months seems realistic. This dramatic increase in the amount of genomic information will have a tremendous impact on biomedical research and on the way that medicine is practiced. When all the human genes are truly known, scientists will have produced a Periodic Table of Life, containing the complete list and structure of all genes and providing us with a collection of high-precision tools with which to study the details of human development and disease. New technologies will facilitate analyses of individual variations in the whole genome and the expression profiles of all genes in all cell types and tissues. The way will thus be paved for systems biology and for deciphering the genetic repertoires of many organisms. The complete genome sequence of humans and of many other species provides a new

L. Peltonen is in the Department of Human Genetics, University of California Los Angeles School of Medicine, Los Angeles, CA 90095–7088, USA. E-mail: lpeltonen@mednet.ucla.edu V. A. McKusick is in the Department of Medical Genetics, Johns Hopkins Medical Institutions, Baltimore, MD 21205, USA.

starting point for understanding our basic genetic makeup and how variations in our genetic instructions result in disease.

The pace of the molecular dissection of human disease can be measured by looking at the catalog of human genes and genetic disorders identified so far in *Mendelian Inheritance in Man* (1) and in OMIM, its online version, which is updated daily (www.ncbi.nlm.nih.gov/omim). For 1100

challenges and opportunities facing members of the research community as they continue to dissect the molecular basis of human disease.

Monitoring Variations in the Genome

Initial analyses of the completed chromosomal sequences suggest that the number of human genes is lower than expected (about 35,000) (2). These findings are

consistent with the idea that variations in gene regulation and the splicing of gene transcripts explain how one protein can have distinct functions in different types of tissue. It also seems likely that obvious deleterious mutations in the coding sequences of genes are responsible for only a fraction of the differences in disease susceptibility between individuals, and that sequence variants affecting gene splicing and regulation must play an important part in determining disease susceptibility. As only a small

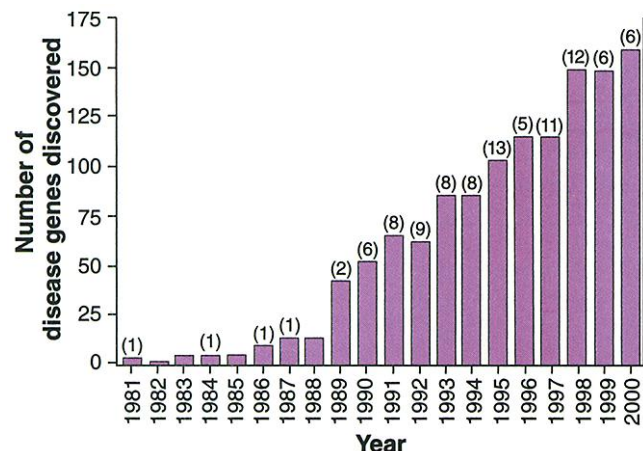
abundance of SNP databases (for example, www.ncbi.nlm.nih.gov/SNP; <http://snp.cshl.org>; <http://hgbase.cgr.ki.se>). However, the follow-up work of understanding how these and other genetic variations regulate the phenotypes (visual characteristics) of human cells, tissues, and organs may well occupy biomedical researchers for all of the 21st century.

Selecting strategies for monitoring the DNA variations associated with human disease requires careful consideration and new innovative methodologies. First, the cost of detecting DNA variations is still too high to enable screening for tens of thousands of SNPs in massive epidemiological study samples. Second, the annotation and cataloging of variations and their frequencies in various populations is not systematically organized. Third, the selection of relevant variants for epidemiological and functional studies is still a guessing game. We know amazingly little about the relative importance of variations in the regulatory and intronic sequences of human genes and how they differ between populations. Fourth, quantitative analyses of the effects of thousands of DNA variations and the "genome-wide" variant profiles that predispose individuals to complex diseases are still in their infancy. All of these issues require methodological developments, coordinated efforts, and better solutions than those currently available to genetic epidemiologists.

Genome-Wide Screening of Gene Transcript Variations

Oligonucleotide and cDNA microarrays have revolutionized the study of differential gene expression in different cells and tissues, enabling numerous scientists to analyze disease processes such as the development of tumors. Microarray techniques are sensitive enough to detect expression of a target gene among 50,000 to 300,000 transcripts (3). The possibility of simultaneously monitoring the transcription rate of virtually every gene from even modest amounts of tissue is now on the horizon. Similar array techniques are also being developed to analyze proteins and their variants (4).

The first reports of genome-wide profiling of gene transcription in tissue from patients with clonal diseases such as



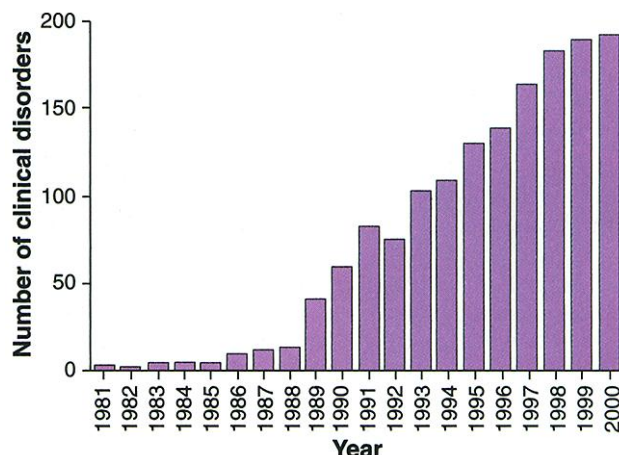
Pace of disease gene discovery (1981 to 2000). The number of disease genes discovered so far is 1112. This number does not include at least 94 disease-related genes identified as translocation gene-fusion partners in neoplastic disorders. Numbers in parentheses indicate disease-related genes that are polymorphisms ("susceptibility genes").

genes, at least one disease-related mutation has been identified (see the figure at the top of this page). Because different mutations in the same gene often result in more or less distinct disorders, the total number of diseases for which OMIM lists mutations approaches 1500 (see the figure at the bottom of this page).

Beginning in 1986, map-based gene discovery (positional cloning) became the leading method for elucidating the molecular basis of genetic disease. Almost all medical specialties have used this approach to identify the genetic causes of some of the most puzzling human disorders. Positional cloning has also been used reasonably successfully in the study of common diseases with multiple causes (so-called complex disorders), such as type I diabetes mellitus and asthma. With the availability of the human genome sequence and those of an increasing number of other species, sequence-based gene discovery is complementing and will eventually replace map-based gene discovery.

With the anatomy of the human genome at hand, the biomedical research community is facing sweeping changes in its methods and strategies (see the table on the next page). Here, we address the major

challenges and opportunities facing members of the research community as they continue to dissect the molecular basis of human disease. The success of global efforts to identify and annotate sequence variations in the human genome—which are called single-nucleotide polymorphisms (SNPs)—is reflected in the



Molecular characterization of clinical disorders (1981 to 2000). The number of clinical diseases characterized (1430) does not include the many neoplastic disorders caused by translocation-related fusion genes.

leukemia (5, 6) show the power of this strategy. But, because insufficient information exists about intra-individual and inter-individual variations in gene expression, profiling human traits beyond those associated with clonal disease processes is still a risky endeavor. Changes in transcription profiles with progression through the cell cycle and during tissue differentiation, as well as variations in expression profiles between individuals, create background noise. This disturbs the detection of "real" signals signifying actual disease-related changes in gene expression. Systematic data collection from expression array experiments and dutiful storage of images in accessible databases would greatly facilitate efforts to understand changes in the transcription of individual genes during various disease processes. If we do not start to carefully catalog collected data now, we will lose a great opportunity to reanalyze cumulative data with the computational and mathematical tools that will be developed in years to come.

Detecting New Metabolic Disease Pathways

The majority of publications reporting genetic studies of complex diseases investigate candidate genes and known metabolic pathways. The major problem with any strategy for analyzing a candidate gene or metabolic pathway is that we look wherever we can (that is, among the candidates that we already know) and we most probably overlook other essential genes or pathways because of our ignorance of human biology. Proteins, the products of genes, perform their functions by interacting with each other in coordinated networks. But, only a fraction of these networks have been identified and characterized through classical biochemistry, structural analyses, and assays of activity. A much more comprehensive view of how proteins interact with each other inside cells will become possible with sequence information from multiple species, including humans. Knowing the full complement of our genes, we should be able to identify all of the metabolic pathways in the human body, no matter how short the half-life of the participating proteins or how small the develop-

mental window during which the pathway is turned on.

Several biocomputing-based strategies have been introduced to construct genetic (and ultimately protein) networks (7). Homology searches among genomes of various species can identify orthologs (closely related gene sequences between species) and paralogs (closely related gene sequences within a species) that encode proteins of known function. Furthermore, shared regulatory motifs and coordinated expression patterns are good indicators of genes that operate in networks. The current search for co-regulated genes is propelled by two very different technologies: biocomputing-based motif searches and expression microarrays. Motif identification combines sequence comparisons and database searches to seek out common promoter and enhancer regions. Expression microarrays directly examine gene co-regulation under a variety of experimental conditions.

Phylogenetic profiling looks for genes that have the same pattern of presence or absence across multiple species. Functionally related genes, which encode proteins that interact with each other, should be subject to similar evolutionary pressures (8). The Rosetta Stone (or fusion domain) method identifies proteins that are sepa-

predicted for the *DAX1*, *SOX9*, and *SRY* genes are very similar (11). These strategies need to be developed further because absolute reliability has not been achieved with any of them (although admittedly some of this unreliability is due to still-missing information). As increasing numbers of eukaryotic genome sequences are deposited in databases and more three-dimensional protein structures are determined from different species, even currently available strategies will yield more reliable information.

Model organisms provide an alternative and complementary approach to the identification of new metabolic pathways. The sophisticated genetic technology with which we study *Drosophila* has provided us with an extraordinary opportunity for analyzing the functions of all fly genes. For example, isolating mutations in fly orthologs of human genes has revealed new tyrosine kinase pathways in *Drosophila*. An oncogenic form of the signaling molecule Ras gives rise to profound developmental abnormalities in the fly eye. By crossing flies that display these abnormalities with heavily mutagenized flies, one can screen for mutations in other genes that modify (either enhance or suppress) the severity of the eye phenotype. This type of "modifier screen" has led to a nearly complete dissection of signaling pathways downstream of Ras, pathways that are conserved in human cells (12). Recently, similar strategies have been applied to the study of the fly insulin-signaling pathway (13). *Drosophila* and other experimental species will be valuable for identifying genes involved in new biochemical pathways, including those implicated in human disease.

Phenotypic Variations in Simplex Diseases

There are major problems associated with dissecting the molecular basis of even simple monogenic diseases caused by mutations in a single gene. Principal among these are the modifying effects of other genes. No gene operates in a vacuum; rather, each gene busily interacts either directly or through its protein product with many other genes and gene products. This results in marked variations in the symptoms of patients with the same disease. Of the 1500 or so monogenic diseases for which the mutated gene has been identified, there are only a few where the effects of other genes on disease pathogenesis have been studied. Existing information about monogenic diseases, such as cystic fibrosis and Hirschsprung disease, demonstrates that certain modifier genes cause

PARADIGM SHIFTS IN BIOMEDICAL RESEARCH

| | | |
|------------------------------|---|---|
| Structural genomics | → | Functional genomics |
| Genomics | → | Proteomics |
| Map-based gene discovery | → | Sequence-based gene discovery |
| Monogenic disorders | → | Multifactorial disorders |
| Specific DNA diagnosis | → | Monitoring of susceptibility |
| Analysis of one gene | → | Analysis of multiple genes in gene families, pathways, or systems |
| Gene action | → | Gene regulation |
| Etiology (specific mutation) | → | Pathogenesis (mechanism) |
| One species | → | Several species |

rate molecules in one organism but fused together in another (9). The gene neighbor method identifies genes that are clustered together on the same chromosome in different species (10).

These approaches already have identified a number of protein networks. For example, recent analysis of sex determination genes from a number of different primate species has provided evidence that, as anticipated, the evolutionary trees

variations in the clinical phenotype of these disorders [for a review, see (14)]. It is probable that many diseases that are considered monogenic will turn out to be "complex" disorders. Such complexity may be attributable to the typically unpredictable effects of gene mutations on the encoded protein and on the metabolic pathway in which the protein acts (15). As exemplified by Marfan syndrome, there could be an expression threshold below which the mutant protein does not cause a disease phenotype (16). In the hemoglobinopathies, for example, modifier genes play an important part in the appearance of clinical symptoms (17). Finally, the recognition that metabolic pathways are only rarely controlled by single rate-limiting steps greatly complicates the prediction of which symptoms a patient will develop based on the gene that is mutated (18, 19). These possibilities all require further analysis in human patients as well as in model organisms before we can understand why the severity of monogenic diseases varies not only with different mutations in the same gene, but also among affected individuals within the same family.

The Genetic Background of Complex Disorders

If there is a challenge to identifying the genes involved in so-called monogenic diseases, clearly this challenge will be far greater for oligo- and polygenic disorders, which have multiple causes. Many of the diagnostic features of these complex diseases—called quantitative trait locus (QTL) disorders—are probably regulated by at least several genes (see the figure on this page).

Well-established mouse models of disease will be crucial for dissecting the molecular basis of complex disorders. Despite millions of years of evolutionary separation, there is close homology between many mouse and human gene sequences and many extended chromosomal regions that have maintained the same genes in the same order. Data sets of this mouse-human synteny are presented in the major human and mouse databases and are becoming even more comprehensive as sequencing of the mouse genome advances (www.ncbi.nlm.nih.gov/Homology; www.informatics.jax.org/menu/homology_menu.shtml).

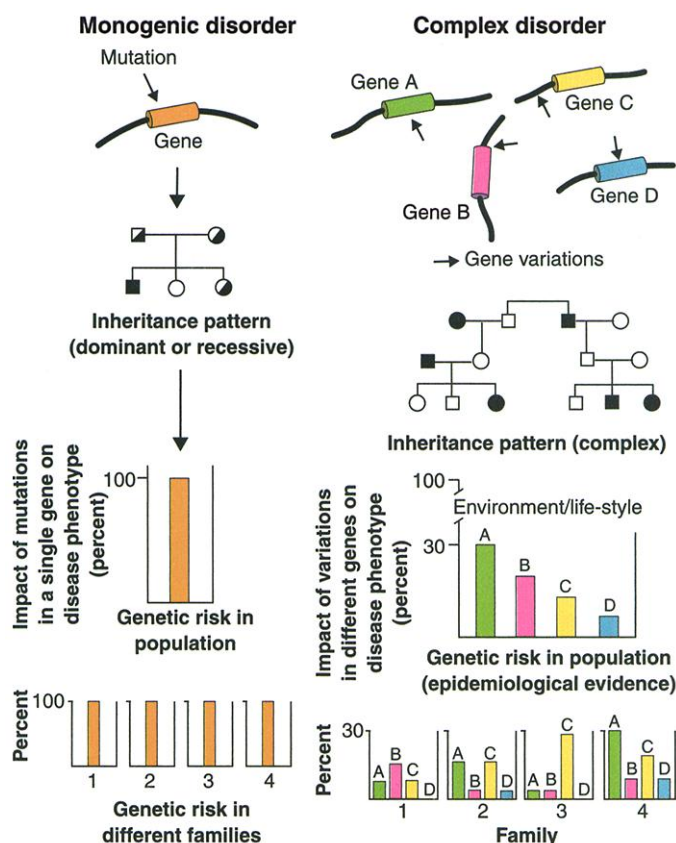
Mouse models, with their short generation times and high breeding efficiency, are extremely useful for

unraveling the causes of complex disorders. They provide shortcuts to disease gene identification, unequivocal proof that a mutation in that gene causes the disease, and rapid dissection of the molecular pathway in which the mutant protein acts. Many human diseases with complex genetic backgrounds have counterparts in the mouse (and in other mammalian species). In a number of instances, the conserved syntenic region that harbors a disease or QTL trait has been identified. The mapping (positional cloning) of a mouse disease gene is relatively straightforward because of the possibility of controlled breeding and crossing of the mice. Suspected chromosomal regions can

be genetically restricted to a section that is small enough to be sequenced and then the mutated gene within this section can be identified. The mutated gene can be further analyzed for functional variations in human study samples (20, 21). The ability to produce transgenic and knockout mutant mice and the possibility of creating congenic strains (in which a limited region of one chromosome from one strain is operating on the background of another strain) removes many of the problems associated with human disease studies that depend on human populations where the genetic background is extremely heterogeneous. In these mouse strains, the functional consequences of mutated genes or even of mutated gene clusters can be studied with high precision.

Similarly, *Drosophila* has served as a valuable model for analyses of normal and aberrant development because the experimental possibilities with this species reach far beyond those of mammals.

One hundred years of classical genetics has provided us with an unprecedented collection of mutant genes at numerous loci that affect fly development and behavior. The catalog of transposon insertions leading to well-described fly mutant phenotypes for a number of genes is available from various databases (www.flybase.org). Because the insertion sites of these transposons have been mapped at the nucleotide level, it is often trivial to go from a human gene to a fly gene and then to the fly mutant on the computer (22). Genes associated with human disease can be investigated in *Drosophila* by searching for the fly homolog, or making flies with P-element insertions, or characterizing previously identified mutations and their associated phenotypes. Furthermore, as with the mouse genome, targeted mutations can be introduced into the fly genome by homologous recombination (23). This provides an extraordinary opportunity to assess the biological ef-



Inheritance of monogenic and complex (multifactorial) disorders. In monogenic diseases, mutations in a single gene are both necessary and sufficient to produce the clinical phenotype and to cause the disease. The impact of the gene on genetic risk for the disease is the same in all families. In complex disorders with multiple causes, variations in a number of genes encoding different proteins result in a genetic predisposition to a clinical phenotype. Pedigrees reveal no Mendelian inheritance pattern, and gene mutations are often neither sufficient nor necessary to explain the disease phenotype. Environment and life-style are major contributors to the pathogenesis of complex diseases. In a given population, epidemiological studies expose the relative impact of individual genes on the disease phenotype. However, between families the impact of these same genes might be totally different. In one family, a rare gene C (Family 3) might have a large impact on genetic predisposition to a disease. However, because of its rarity in the general population, the overall population effect of this gene would be small. Some genes that predispose individuals to disease might have minuscule effects in some families (gene D, Family 3).

fects of disease genes in *Drosophila*.

Mouse and fly, however, are not ideal study subjects for all complex diseases. But, so far, the lack of genomic resources for nonhuman primates has limited the possibilities for genomic and genetic comparisons between humans and the great apes. The value of such comparisons for biomedical research is increasingly apparent. The benefits will be particularly profound for elucidating the genes involved in human behavior. Opportunities for comparing the human genome with those of other primates have opened up with the development of new genomic resources, such as the recently completed whole-genome marker map for the baboon (24). Sequencing of the chimpanzee genome should be a high priority because such a resource would be enormously beneficial for understanding those diseases that cannot be studied in lower species. We should also learn a few lessons from nonanimal species. For example, the successful positional cloning of a single gene responsible for the QTL of tomato fruit size was recently reported (25). Some human QTLs may also prove to have a simple basis, with a single gene responsible for the complex disorder.

Although experimental species are of great value for the initial identification and functional analysis of complex disease genes, final evidence for the involvement of these genes in human diseases must come from extensive epidemiological studies, preferably in different populations. Statistical methods that analyze the multiple variations and quantitative trait components of complex diseases are being developed (26). There are still too many gene loci and not enough gene mutations implicated in complex disorders. The debate over gene identification strategies continues—Which populations should be studied? What are the optimal strategies for statistical analysis that can take advantage of both vast genetic data sets and quantitative phenotype information? How will associations between a genetic marker and a clinical phenotype be validated?

The best strategy will probably be to collect large epidemiological study samples from many different populations. Predictions of the minor effects of multiple genes in complex disorders will require efficient collaboration among research groups, data sharing, and pooled data analyses. Seamless collaborations among clinicians, epidemiologists, geneticists, mathematicians, and computer experts will be needed to solve the genetic underpinnings of complex diseases that affect the lives of millions.

Dissecting Interactions Between Genes and Environment

Most common human diseases represent the culmination of lifelong interactions between our genome and the environment. Predicting the contribution of genes to complex disorders is still a challenge, and determining the interactions between genes and the environment during any disease process is a daunting task. Many human diseases, such as hypertension, coronary artery disease, and even some psychiatric disorders, represent quantitative traits that are caused by interactions among genes and between genes and the environment. For instance, QTL genes that contribute to elevated lipid levels in the blood may only be expressed if the individual eats a high-fat diet. Epidemiological study cohorts that carefully report and register environmental factors—such as smoking, type of diet, exercise habits, events during fetal life and early childhood (for example, infections)—will be of immense importance when combined with genetic risk profiling. If these variables are measured and carefully registered, then not only can interactions between these variables be monitored, they may even be included as covariants in QTL-type analyses. Incorporation of suspected or known large-scale interactions among genes will require new analytical strategies. The functional importance of identified DNA variations can be established by taking advantage of experimental opportunities in other species—for example, by introducing specific DNA variations into animal strains that have a well-defined genetic background or that live in an environment that can be precisely controlled.

Genetic Information in Health Care

We are rapidly advancing upon the post-genomic era in which genetic information will have to be examined in multiple health care situations throughout the lives of individuals. Currently, newborn babies can be screened for treatable genetic diseases such as phenylketonuria. Perhaps in the not-so-distant future, children at high risk for coronary artery disease will be identified and treated to prevent changes in their vascular walls during adulthood. Parents will have the option to be told their carrier status for many recessive diseases before they decide to start a family. For middle-aged and older populations, we will be able to determine risk profiles for numerous late-onset diseases, preferably before the appearance of symptoms, which at least could be partly prevented through dietary or pharmaceutical interventions. In the near future, the monitoring of individual drug response profiles

with DNA tests throughout life will be standard practice (27). Soon, genetic testing will comprise a wide spectrum of different analyses with a host of consequences for individuals and their families—an issue worth emphasizing when explaining genetic testing to the public.

The challenge for health care professionals will be to correctly interpret the outcome of genetic testing for their patients, their patients' families, and for society in general. Genetic counselors, who explain the purpose and results of genetic tests, will be crucial for helping individuals to make informed decisions, particularly when test results indicate the possibility of disease. Current training programs, including those in medical schools, do not adequately teach students how to deal with these challenges.

The tremendous potential for efficient information transfer via the Internet can and should be used to inform the public of the possibilities provided by the genomics era. However, when it comes to sensitive and very personal aspects of genetic information, traditional contact with health care professionals is still the most appropriate route. Reaping the fruits of the human genome sequencing project through alleviating the suffering of patients will only be possible if available genetic information is combined with the skilled professionalism of health care workers and ethically solid standards.

References

1. V. A. McKusick, *Mendelian Inheritance in Man* (Johns Hopkins Univ. Press, Baltimore, ed. 12, 1998).
2. I. Dunham et al., *Nature* **402**, 489 (1999).
3. D. J. Lockhart, E. Winzler, *Nature* **405**, 827 (2000).
4. A. Pandey, M. Mann, *Nature* **405**, 837 (2000).
5. T. R. Golub et al., *Science* **286**, 531 (1999).
6. C. M. Perou et al., *Nature* **406**, 747 (2000).
7. D. Eisenberg, E. M. Marcotte, I. Xenarios, T. O. Yeates, *Nature* **405**, 823 (2000).
8. M. Pellegrini et al., *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4285 (1999).
9. E. M. Marcotte et al., *Nature* **402**, 83 (1999).
10. R. Overbeek et al., *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2896 (1999).
11. M. Patel et al., *Am. J. Hum. Genet.* **68**, 275 (2001).
12. F. D. Karim et al., *Genetics* **143**, 315 (1996).
13. R. Bohni et al., *Cell* **97**, 865 (1999).
14. K. M. Dipple, E. R. B. McCabe, *Mol. Genet. Metab.* **71**, 43 (2000).
15. ———, *Am. J. Hum. Genet.* **66**, 1729 (2000).
16. F. Ramirez, B. Gayraud, L. Pereira, *Ann. Med.* **31**, 202 (1999).
17. Y. P. Chang et al., *Br. J. Haematol.* **96**, 806 (1997).
18. S. Krauss, P. A. Quant, *J. Theor. Biol.* **182**, 381 (1996).
19. M. D. Brand, *J. Exp. Biol.* **200**, 193 (1997).
20. M. Stoll et al., *Genome Res.* **10**, 473 (2000).
21. J. Lusis, *Nature* **407**, 233 (2000).
22. G. M. Rubin et al., *Science* **287**, 2204 (2000).
23. Y. S. Rong, K. H. Golic, *Science* **288**, 2013 (2000).
24. J. Rogers et al., *Genomics* **67**, 237 (2000).
25. A. N. Frary et al., *Science* **289**, 85 (2000).
26. J. Blangero, L. Almasy, *Genet. Epidemiol.* **14**, 959 (1997).
27. A. Roses, *Nature* **405**, 857 (2000).