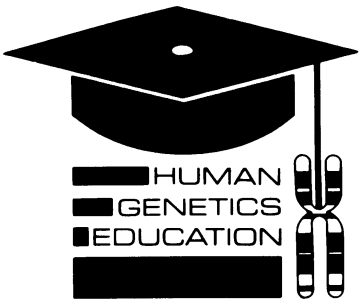
Am. J. Hum. Genet. 54:129–136, 1994

The Evolving Genome Project: Current and Future Impact

Eric P. Hoffman

Department of Molecular Genetics and Biochemistry, Department of Human Genetics, and Department of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh



Summary

The National Institutes of Health/Department of Energy Human Genome Project has been funding directed research for only 5 years, and it is understandably difficult to cite important research advances directly attributable to the project. However, the project has been constructive in fostering multidisciplinary group research and an inspiring and synergistic "just do it" attitude in both political and scientific circles, domestically and abroad. This collaborative spirit has spawned large-scale genetic and physical mapping projects, with the most impressive and useful results to date being the dense genetic maps produced by the Généthon, a French organization largely supported by the French muscular dystrophy association. With the genetic and physical map reagents now becoming available, disease-gene cloning is proceeding at an increasingly rapid pace. More important than the predictable acceleration of disease-gene mapping are the unpredictable benefits: Will a dense PCR-based dinucleotide-repeat genetic map open novel alternative approaches to disease-gene isolation? Will it become possible to localize disease genes by simply analyzing unrelated, isolated probands rather than the rarer "extended family"? Proband-based "linkage-disequilibrium cloning" may become possible if adequate density, informativeness, and stability of polymorphic loci are obtained. In addition, "genome exclusion cloning" will be added to the established positional, candidate-gene, and functional-disease-gene-cloning experimental approaches. The anticipated exponential expansion of human genetic disease information over the remainder of the 10-year tenure of the Human Genome Project unveils critical yet unresolved issues for medical education and the practice of medicine. As we strive for the epitome of preventive medicine—a personal genetic propensity database provided at birth—medical education must tool up to teach the meaning and use of this valuable information. The insurance industry seems ill-equipped to use this information. Will the Human Genome Project unintentionally force the hand of nationalized health care?

Introduction: Long-Term Objectives, Past and Present

A thorough discussion of the history of the Human Genome Project can be found in an accompanying paper and elsewhere (Watson 1990; Jordan 1992; Juengst 1994). In brief, government committees recommended the initiation of a Human Genome Project in 1988. The Department of Energy (DOE) and National Institutes of Health (NIH) began earmarking funds for directed research in 1987 and 1988, respectively, and the National Center for Human Genome Research division of the NIH was established in 1989. The research directives began with an initial 5-year (1990–95) plan based on the combined $200-million/year funding of the DOE and NIH. The strategies for the remaining 10 years of the total 15-year $3-billion Human Genome Project were to be designed as evolving technology and goals were reassessed.

About 4 years of the initial 5-year period have elapsed. During this time, dozens of review articles have been published proclaiming a future profound impact of the Human Genome Project on medicine (Green and Waterston 1991; Caskey and Rossiter 1992; Richardson 1992; Rosenburg 1992). However, the quickly evolving nature of research has made specific predictions difficult. Two concrete and partially realized goals of the current 5-year period have been the devel-

Received February 8, 1993; revision received June 9, 1993.

Address for correspondence: E. P. Hoffman, BST W1211, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261.

0002-9297/94/5401-0020$02.00

opment of informatics (Erickson 1992; Pearson, submitted) and the establishment of a research program in ethics (Annas 1992; Garver and Garver 1994; Juengst 1994; Parker 1994). Recent developments in both genetic and physical mapping have begun to crystalize perceptions of the future impact of the project. This discussion will begin with implications of Human Genome Project data for health care and medical education and will then describe the research and research-policy trends on which these projections are based, using the four research directives originally set forth by the organizers of the project: genetic maps, physical maps, sequencing, and model organisms.

Implications for Health Care

Rapidly accumulating data from "human genetics" research, facilitated and expedited by, but not synonymous with, the "genome project," will eventually lead us to an intimate knowledge of the propensities of each individual for multifactorial disorders. The day of the personal DNA profile provided at birth, complete with calculated risks of various cancers (e.g., breast or colon), heart disease, alcoholism, and many other conditions, could be an actuality by the time current first-year medical students begin to practice medicine. The potential impact on length and quality of life, and on reduced cost of health care, are tremendous. For example, a private-practice neurologist in a Washington, DC, suburb currently prescribes ~$5 million in magnetic resonance imaging (MRI) scans per year, primarily to rule out the small possibility of a vascular aneurism or tumor as the cause of a headache. He claims he is typical. This neurologist and thousands of other physicians across the country are forced into the costly "rule-out" situation by threat of malpractice, the patient's knowledge of the wonders of expensive procedures such as MRI, and ready reimbursement by insurance companies. Everyone realizes this scenario must change, but who is to determine who gets the MRI and who does not? Now picture the situation 5 years from now, when each person's genetic risk for cerebral aneurism or brain tumor can be determined. These risks can be factored into logical and rational cost/benefit calculations acceptable to patients, physicians, insurance companies, and lawyers alike. A similar example, already in practice, concerns the recent identification of a colon cancer susceptibility gene and the implementation of colonoscopy screening for malignant growths in patients genetically determined to be at high risk (Aaltonen et al. 1993; Peltomaki et al. 1993).

A second impending scenario is the extension of neonatal screening programs beyond identification of curable metabolic disorders, to include molecular testing for genetic propensities. Learning of a newborn's propensity for heart disease, the family and physician can plan and execute a program of diet and exercise that will dramatically increase the life expectancy of the child and that will reduce the requirement for multiple surgical interventions 50 years later—an ideal program of preventive health care. The accurate knowledge of risks, attainable through molecular studies, could have a powerful influence on lifestyle, with consequent savings on health care that are potentially substantial.

These risk calculations, which should dramatically benefit patient and doctor, are the same numbers sought by insurance company actuaries. Why not charge premiums for identified health risks? The tensions among insurance company, physician, and patient are already obvious: Lisch nodules detected incidentally by eye exam are not reported by the ophthalmologist for fear that the stigmatizing yet benign diagnosis of neurofibromatosis may threaten the patient's insurance coverage. With the majority of health care in the United States still focused on treatment rather than on prevention, the predictive nature of genetic information is of little positive value to the insurance industry. Thus, human genetics may force the hand of socialized medicine in the United States.

Another foreseeable shortcoming of the current health-care system pertains to the interpretation of molecular data by health-care professionals. In general, the level of knowledge of primary-care physicians regarding molecular diagnostic information is not dissimilar to that of the patients being counseled. Many medical schools are now altering curricula to increase the amount of medical genetics taught—the University of Pittsburgh Human Molecular Genetics course includes 26 hours of lectures and over 10 hours of group discussion focusing on issues including insurance, ethics, molecular diagnostics, and gene therapy. Nevertheless, such attempts to prepare future physicians for the impending impact of medical genetics are probably insufficient and are often ill-received by the students, particularly those with weaker undergraduate science backgrounds, who find the genetics thought process obscure and irrelevant. Perhaps one solution to this problem would be the requirement of mandatory rotations through medical genetics. Currently, only a minority of medical schools offer any type of medical genetics clinical rotations, and those that do, offer it as an elective.

Genetic counseling–training programs and their graduates have admirably filled the gap between current knowledge and patient care. However, even these programs have trouble keeping pace with molecular genetic developments, as well as with the demand for new counselors. The very limited molecular genetic knowledge base of physicians now entering practice dictates that the use of genetic counselors must expand beyond the major academic medical centers currently housing them.

Directive 1: A Genetic Map of the Human Genome

A genetic map consists of polymorphic markers spaced at determined intervals. The primary utility of a genetic map is to facilitate the localization of disease genes by family-based genetic linkage studies. Genetic maps of humans have been available for decades: knowing the segregation pattern of two phenotypic or biochemical markers qualifies as a "genetic map" of sorts. The goal of the Human Genome Project is to make this map "complete and fully connected." Clearly, "complete" and "connected" are relative terms, subject to interpretation. The stated 5-year goal of "markers spaced an average of 2–5 cM apart" may sound ideal, yet such a map would be useless if the constituent markers showed a low degree of heterozygosity, were expensive or labor-intensive to use, or were difficult to interpret.

The recent employment of simple sequence repeats detected by PCR (e.g., CA repeats and microsatellites) as the polymorphic markers of choice (Weber and May 1989) has provided the technological advance for which the designers of the Human Genome Project had hoped. These repeats are easy to find (occurring, on average, every 30,000 bp throughout the genome), generally show a very high PIC (meioses are often >80% informative), and are inexpensive, fast, and technically simple to use. Indeed, even acquisition of the single specific reagent required for each locus has become streamlined; over the past year, >1,000 CA repeat PCR primer pairs have become available through private companies for only 20¢/reaction, and the number of primers will probably double over the next year.

Who has isolated all the CA repeats, and what density have they reached? CA-repeat loci have been accumulating on three fronts: "directed" isolation around genes of interest, isolation by chromosome-specific university-based groups funded by the NIH Genome Institute, and large-scale genomewide isolation by a few research groups. The disadvantage of the directed and chromosome-specific strategies is that the resulting genetic maps can be quite uneven, with some regions of the genome replete with dozens of markers while other regions remain barren. The genomewide isolation strategy yields genetic maps of uniform density and recently has proved its utility: Weissenbach et al. (1992) have provided the Genome Database with the sequences for 808 CA-repeat PCR primer pairs distributed throughout the genome. The majority (80%) of these loci show heterozygosities of >70% and have an average spacing of ~5 cM. This landmark work attained the genetic mapping goal of the first 5-year period of the NIH/DOE Human Genome Project, 2.5 years ahead of schedule.

Should we applaud the NIH, the DOE, and the Human Genome Project for this second-generation genetic linkage map? Actually, the advent of simple sequence repeats largely preceded the initiation of directed Human Genome Project funding. Moreover, the pivotal work of Weissenbach et al. was done in a French institute (Généthon) established and funded by the French muscular dystrophy association (AFM), using monies acquired through nonprofit, private fundraising (telethons).

It is important to ask why the French were able to surpass the American research machine, with respect to both genetic and physical (described below) mapping, despite the targeted domestic funding of the Human Genome Project. The approaches were fundamentally different: the French opted for a single group working on genetic mapping of the entire genome, in a very intensive manner, while the NIH Genome Institute originally opted for different university-based groups each working on a single chromosome. The logic underlying the chromosome-based strategy was that a multifaceted, multiuniversity approach would identify the optimal experimental strategies, rather than placing all eggs in one basket, and, further, that the funding of several groups would spread the money and technology around, preventing the creation of a single specialized center. At the same time, scientists with an interest in a specific chromosome (an interest primarily due to disease localization) could feel comfortable focusing their efforts on that chromosome. In effect, NIH and DOE created a number of small companies dedicated to developing alternative technologies for a product (a genetic and physical map), hoping to identify the most successful strategies in the process. Two years ago, it looked as if the American NIH Genome Institute had made the better choice: the French Généthon was ini-

tially set up for large-scale automated Southern blot analysis of RFLPs, only to have the methodology become obsolete with the advent of PCR-based dinucleotide repeats. However, the French aggressively retooled for PCR-based dinucleotide analyses, making impressive strides in the end analysis.

The NIH Genome Institute has recently decided to fund a similar genomewide marker-isolation project (directed by Jeffrey Murray at the University of Iowa), for $14 million over 4 years (Roberts 1992). Between the combined efforts of the French and American groups, it is anticipated that the density of the genetic map may approach 0.5 cM (~1 marker/500,000 bp) within the next few years. It is important that the change in the funding strategy of the NIH Genome Institute has come at the expense of finding university homes for about half of the chromosomes: the chromosome-based-center grant philosophy is apparently being phased out.

Is the genetic map complete? The current density of PCR-based VNTRs has, on average, already reached the 5-cM goal, which was to be achieved by 1995. However, there is variability in the density: many gaps of 20 cM still exist, and it seems that CA repeats are difficult to isolate near telomeres, leading to uncharted regions. As the holes in the maps become a major hurdle, the chromosome-based strategy may again become the more viable approach (just as NIH is phasing it out).

While the use of CA repeats permits acquisition of high-density and informative genetic maps very quickly, it also comes with some unwanted baggage. First, the results can be difficult to interpret: there are often dozens of alleles for any given locus, alleles often differ by only 2 nt in length (a 1% change in size), and other features of the PCR products cause the patterns to be complex. Thus, mistypings of alleles are a recognized technical problem. Even a single mistyping in a genetic linkage study can ravage the accuracy of its conclusions. Today, most laboratories search for the statistical "outliers" (meiotic events that do not fit with the rest of the analysis) as signposts for typing errors. However, this approach can be likened to Russian roulette in the statistical game of genetic linkage studies. An alternative approach that is being pursued in our laboratory and in others is automated data acquisition and computerized interpretation of CA repeat haplotypes generated on automated sequencers.

Another potential problem with CA repeats is their relatively high mutation rate. In the recent Weissenbach et al. (1992) data set, the mutation rate of the 814 markers studied was 0.1%, which is an order of magnitude higher than that for the mutation-prone Duchenne muscular dystrophy (DMD) gene. This mutability could cause problems in family linkage studies, particularly if an allele being followed is converted to a different allele already segregating in the same family, causing misinterpretation of the linkage data. Nevertheless, the 1/1,000 meioses mutation rate should pose such a problem only occasionally (for <1 family/100 studied).

What is the practical benefit to the human genetics community of the rapidly evolving genetic map? Given a "nice" monogenic disease (fully penetrant with genetic homogeneity and large families available), a single graduate student can map the guilty gene within 2 years. This is 1–2 orders of magnitude less effort than has characterized most genome-based genetic mapping studies to date. Clearly, this is of immediate benefit to the human genetics community: disease-gene mapping, with the corollary disease-gene isolation and molecular diagnostics of segregating families, will proceed at breakneck speed.

Many of the most common and most insidious human health problems are recognized as multifactorial, with an underlying genetic component termed "propensity." The genetic mapping and identification of genes imparting genetic propensity for a disease permits the subsequent identification of the nongenetic factors that contribute to the development of that disease: a group of people with identical genetic propensity can be studied longitudinally to determine environmental (extragenic) risk factors. The identification of risk factors for heart disease, cancer, glaucoma, obesity, susceptibility to debilitating infectious disease, and other common disorders would enable truly directed and rational "preventive medicine": a patient would simply be advised of his risk factors and would be counseled accordingly.

The successful mapping of complex genetic disorders will probably proceed through one of two methods: (1) linkage studies within large disease-segregating families or (2) linkage-disequilibrium studies using large numbers of unrelated affected individuals. Family linkage studies are often limited by the availability of large families, genetic homogeneity, or both and also by the correct modeling of disease inheritance (Kelsoe et al. 1989; Pauls 1993). Isolation of disease genes (both simple and complex) by proband-based linkage disequilibrium may offer an alternative approach to family studies. Many monogenic disorders show a single mutation that is responsible for the majority of cases (e.g., sickle

cell anemia, medium-chain acyl-dehydrogenase, Tay-Sachs disease, and cystic fibrosis [CF]). All such disorders show a founder effect, where a single mutation becomes embedded in the gene pool at a relatively high frequency for reasons that are generally not clear. These same mutations show linkage disequilibrium with nearby polymorphic loci. The finding of a specific haplotype in linkage disequilibrium with a phenotype has generally *followed* genetic mapping of the disease locus. With dense and informative genetic maps, linkage disequilibrium could conceivably be used to *identify* disease loci: one would simply study large numbers of polymorphisms in isolated probands with a well-defined disease showing a predilection for certain ethnic populations. Certain polymorphisms should show a skewed allele distribution frequency in the patients compared with the control population, thus indicating the location of the culpable gene. The design of such studies should provide a challenge for statistical geneticists in the years to come. One feature that must be built into future statistical modeling of linkage disequilibrium cloning is the relative instability of dinucleotide repeat loci, although this parameter simply broadens the concept of "allele" to include more than one dinucleotide repeat. For example, a TA repeat found in the CF gene shows a "clustering" of alleles (30–32 repeats) in strong linkage disequilibrium with delta-F508, while non–delta-F508 mutations and normal alleles show a much different pattern of allele frequencies over the possible 7–56 repeats (Zielenski et al. 1991).

Another developing approach is gene mapping through "genomic exclusion." The high density of highly informative markers makes it possible to search for small regions of the genome that are shared by all affected individuals but that are not shared by unaffected individuals. For example, given an X-linked recessive pedigree, it is possible to localize a gene to within ~1 million bp with only 10 meiotic events. A variation of the affected sib pair method (Weeks and Lange 1992; Ward 1993), this experimental design simply constructs concordance maps of regions of the X chromosome. As with many of the methods relying on dense and complex genetic linkage data, the development of computerization of both data collection and interpretation is critical, particularly if the analyses are extended from a small region to the entire genome.

Finally, it should be mentioned that one of the original directives of the Human Genome Project was to have each polymorphic marker identified with a PCR-based "sequence-tagged site" (STS) so that the marker could be placed on the coevolving physical map. The functional replacement of RFLPs with dinucleotide repeat polymorphisms has obviated this objective.

Directive 2: A Physical Map of the Human Genome

The goals for the cloning of the human genome, as stated in the 5-year plan, were twofold: (1) to develop unique sets of PCR primers that would recognize sites ~100,000 bp apart across the entire genome and (2) to begin to generate large segments (2,000,000 bp, equivalent to ~1% of the X chromosome) of contiguous cloned DNA at different defined sites in the genome.

As with genetic mapping, the initial approach taken by the NIH Genome Institute toward this directive was to fund specific groups for specific chromosomes. Thus, "genome centers" were established that focused on one or more chromosomes and whose goal it was to work on the physical maps of those chromosomes as defined by the 5-year plan. Over the past 3 years, approximately half of the human chromosomes have been "assigned" to certain universities in the United States. Again, the French (this time the CEPH) showed that they could develop megabase YAC libraries covering the genome, an accomplishment that, by its nature, cast doubt on the chromosome-specific-center strategy of the NIH. During the past year, the NIH has suddenly shifted policy, awarding nearly $24 million to Eric Lander (Massachusetts Institute of Technology) and collaborators (including Daniel Cohen at CEPH) for a 5-year project to construct physical maps of both humans and mice (Roberts 1992).

This new policy may lead to a complete physical map of the human genome in just a few years, thereby vastly exceeding the expectations of the NIH/DOE 5-year plan. Moreover, the centralization of this research should ensure and facilitate the dissemination of reagents to other investigators. There are also some impending negative effects of this policy shift. Both the DOE and the NIH have now committed a substantial part of their annual budget to a few centers, severely restricting the funds available for new individual research awards and new genome centers. For example, some approved grant applications on Human Genome Project topics are now being transferred from the cash-strapped NIH Genome Institute to other medically oriented NIH institutes for funding. Another evolving variable in the funding policies of the NIH Genome Institute involves the recruitment of Francis Collins as the head of the National Center for Genome Research

at the NIH. Collins hopes to start an intramural research facility for genome research at the NIH Bethesda campus, with funding to reach $40 million by 1995 (representing 40% of the current NIH Genome Institute annual budget) (Thompson 1993). In terms of 1993 numbers, the proposed NIH intramural program and the large Iowa and MIT centers would soak up 35% of NIH funds earmarked for genome research. The currently shifting sands of Human Genome Project funding strategy may remain unstable for some time to come.

The anticipated fulfillment of the complete physical map far ahead of schedule will be a boon to both basic and applied human genetics. When a disease gene is genetically mapped, the corresponding genomic locus in large YACs can be easily obtained from the physical mapping projects, and the culpable disease gene subsequently identified. The identification of the gene (for either monogenic or complex disorders) then leads to mutation identity and, consequently, direct molecular diagnostics. Again, the acceleration of disease-gene isolation cannot be underestimated.

There are a few caveats to this idyllic picture. First, knowledge of the genetic location of a disease gene does not always lead immediately to the identification of that gene. The elusive Huntington disease gene serves as the most dramatic example of this point (Hoffman and Jaffurs 1993; Huntington's Disease Collaborative Group 1993). An emerging objective of the Human Genome Project, isolation and mapping of expressed sequences (cDNAs), could circumvent this problem to some extent: any segment of the physical map could come with a cohort of partially characterized genes located within that segment. Thus, we may begin to see the marriage of the "positional cloning" approach (so successful for DMD, CF, and neurofibromatosis, among many others) with the "candidate-gene" approach (e.g., periodic paralysis, Marfan disease, and retinitis pigmentosa). Second, the recent change of policy at the NIH Genome Institute epitomizes the feared funding of "big science" over smaller individual research grants. The scale-down of huge defense contracts may simply find a publicly acceptable parallel in the buildup of technically demanding and hardware-intensive biomedicine. Whether sufficient funds will be allocated to support both big and small biomedical research must await directives from changing political and scientific administrations. One must feel some compassion for the NIH Genome Institute as its members try to predict optimal experimental strategies, as both politics and science rapidly evolve.

Directive 3: Sequencing of the Human Genome

"Sequencing of the human genome" is considered by many to be the manifesto of the Human Genome Project. To the contrary, sequencing has taken a back seat in the 5-year plan, largely because it is too expensive and would not yield much practical data that could be applied to human health problems. The goals stated in the 5-year plan were to reduce the cost of sequencing to 50¢/nt (current cost $2–$5/nt) and to sequence a total of ~10,000,000 bp at a few sites.

This aspect of the Human Genome Project is currently of least concern to the human genetics community: the data generated to date are technical rather than practical. In the future, it is hoped that the sequence data will facilitate identification of genes in the physical map, thereby providing candidate genes for genetic diseases.

Directive 4: Mice, Worms, Flies, and Yeast

The 5-year plan called for a genetic map of the mouse and for physical maps of worms (*Caenorhabditis elegans*), flies (*Drosophila melanogaster*), and yeast (*Saccharomyces cerevisiae*). Why? The hypothesis is that these "model" organisms serve as paradigms for humans, in both health (basic molecular and cellular physiology) and disease (identification of new disease genes). The genetic mapping of these species was already well underway before the inauguration of the Human Genome Project, and the project probably has not accelerated genetic and physical mapping efforts in these organisms to the same extent that it has in humans.

Are there any past, current, or future benefits to human genetics that are attributable to model organisms? Genetic diseases in animal models are often cited as important tools in the elucidation of homologous human disease. However, more often the genetic basis for human disease is uncovered first, followed by the identification or development of homologous animal models by using the human-derived molecular tools. Chronology of discovery does not dictate utility, as all animal models prove very valuable for understanding human disease pathogenesis and treatment, regardless of whether the human or animal gene was identified first. However, the chronological distinction is relevant to discussion of the Human Genome Project: genetic and physical maps of animal models are generally not necessary if the human disease gene is identified first. Illustrative examples include malignant hyperthermia in pigs (MacLennan and Phillips 1992); mouse, dog, and

cat models of DMD (Hoffman et al. 1987; Cooper et al. 1988; Gaschen et al. 1992); and the horse model of periodic paralysis (Rudolph et al. 1992)—all of which were identified after discovery of the human disease gene. On the other hand, there are a number of striking examples of disorders in mice that, in retrospect, are proving to be the molecular precedent for the homologous human disease—the murine equivalents of juvenile-onset insulin-dependent diabetes (Acha-Orbea and McDevitt 1987) and Charcot-Marie-Tooth peripheral neuropathy (Valentijn et al. 1992), to name two. Thus, it is probably true that disorders of genetically well-mapped mammals (e.g., mice) will continue to serve as paradigms for certain human diseases.

Genetic disorders of *D. melanogaster, Escherichia coli, C. elegans,* or *S. cerevisiae* do not yet, and probably will not, serve as reliable paradigms for human disease: the complex nature of the pathophysiological component of most genetic diseases precludes accurate predictions of cross-species genotype/phenotype relationships. Nonetheless, coding sequences of genes are more easily and quickly identified in these simpler organisms, and the human homologues can then be identified and localized on the human genetic and physical maps. In this context, the lower-species model organism could have value as an indirect source of "candidate genes" for human disorders. Furthermore, basic principles of genetics and biochemistry can be more readily learned in these well-characterized and manipulable species. However, in general, the direct impact of invertebrate-model-organism genome studies on human genetics is not expected to be immediately apparent.

Conclusion

Medical genetics is continuing to carve out a visible and respected place in medical practice, and the Human Genome Project will certainly accelerate this process. The human genetics community will find itself in a position of ever-increasing importance in medicine: as the tools developed by the Human Genome Project initiatives are used by disease-based research projects, the identification, collection, and characterization of families becomes crucial. The input of experienced medical geneticists and counselors is imperative to the current and future family studies that will quickly define both monogenic and complex genetic traits. The role of the human geneticist as educator of both physician and patient will become increasingly important and challenging.

Acknowledgments

The author is grateful to Michael Gorin, Daniel Weeks, W. Gregory Feero, Rita Hoop, and Anu Bhaskar for very helpful critiques of the text and to Mark Perlin for stimulating collaborations. The author's laboratory is supported by National Institutes of Health grants AR41025, NS28403, and NS32084, by the Muscular Dystrophy Association, and by the March of Dimes.

References

Acha-Orbea H, McDevitt HO (1987) The first external domain of the non-obese diabetic mouse class II, I-A beta chain is unique. Proc Natl Acad Sci USA 84:2435–2439

Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Pylkkanen L, Mecklin JP, Jarvinen H, et al (1993) Clues to the pathogenesis of familial colorectal cancer. Science 260:812–816

Annas GJ (1992) The Human Genome Project as social policy: implications for clinical medicine. Bull N Y Acad Med 68:126–134

Caskey CT, Rossiter BJF (1992) The Human Genome Project: purpose and potential. J Pharm Pharmacol 44:198–204

Cooper BJ, Winand NJ, Stedman H, Valentine BA, Hoffman EP, Kunkel LM, Scott MO, et al (1988) The homologue of the Duchenne locus is defective in X-linked muscular dystrophy of dogs. Nature 334:154–156

Erickson D (1992) Hacking the genome. Sci Am (April), pp 128–137

Garver KL, Garver B (1994) The Human Genome Project and Eugenic Concerns. Am J Hum Genet 54:148–158

Gaschen FP, Hoffman EP, Gorospe JRM, Cardinet GH, Ulh EW, Senior DF, Pearce LK (1992) Dystrophin deficiency causes lethal muscle hypertrophy in cats. J Neurol Sci 110:149–159

Green ED, Waterston RH (1991) The Human Genome Project: prospects and implications for clinical medicine. JAMA 266:1966–1975

Hoffman EP, Brown RH, Kunkel LM (1987) Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell 51:919–928

Hoffman EP, Jaffurs D (1993) Huntington's disease: an expanding enigma. Curr Biol 3:456–459

Huntington's Disease Collaborative Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 72:971–983

Jordan E (1992) The Human Genome Project: where did it come from, where is it going? Am J Hum Genet 51:1–6

Juengst E (1994) Human genome research and the public interest: progress notes from an American science policy experiment. Am J Hum Genet 54:121–128

Kelsoe JR, Ginns EI, Egeland JA, Gerhard DS, Goldstein AM, Bale SJ, Pauls DL, et al (1989) Re-evaluation of the linkage relationship between chromosome 11p loci and the gene for bipolar affective disorder in the Old Order Amish. Nature 342:238–243

MacLennan DH, Phillips MS (1992) Malignant hyperthermia. Science 256:789–794

Parker LS (1994) Bioethics for human geneticists: models for reasoning and methods for teaching. Am J Hum Genet 54:137–147

Pauls DL (1993) Behavioural disorders: lessons in linkage. Nature Genet 3:4–5

Peltomaki P, Aaltonen LA, Sistonen P, Pylkkanen L, Mecklin JP, Jarvinen H, Green JS, et al (1993) Genetic mapping of a locus predisposing to human colorectal cancer. Science 260:810–812

Richardson WC (1992) Summary: the Human Genome Project, an agenda for science and society. Bull N Y Acad Sci 68:162–170

Roberts L (1992) NIH takes new tack on gene mapping. Science 258:1573

Rosenburg, LE (1992) The Human Genome Project. Bull N Y Acad Med 68:113–125

Rudolph JA, Spier SJ, Byrns G, Rojas CV, Bernoco D, Hoffman EP (1992) Periodic paralysis in quarter horses: a sodium channel mutation disseminated by selective breeding. Nature Genet 2:144–147

Thompson L (1993) Healy and Collins strike a deal. Science 259:22–23

Valentijn LJ, Baas F, Wolterman RA, Hoogendijk JE, van den Bosch NHA, Zorn I, Gabreels-Festen AAWM, et al (1992) Identical point mutations of PMP-22 in *Trembler-J* mouse and Charcot-Marie-Tooth disease type 1A. Nature Genet 2:288–291

Ward PJ (1993) Some developments on the affected-pedigree-member method of linkage analysis. Am J Hum Genet 52:1200–1215

Watson JD (1990) The Human Genome Project: past, present, and future. Science 248:44–49

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Weeks DE, Lange K (1992) A multilocus extension of the affected-pedigree-member method of linkage analysis. Am J Hum Genet 50:859–868

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, et al (1992) A second-generation linkage map of the human genome. Nature 359:794–801

Zielenski J, Markiewicz D, Rininsland F, Rommens J, Tsui L-C (1991) A cluster of highly polymorphic dinucleotide repeats in intron 17b of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. Am J Hum Genet 49:1256–1262