Appendix B

Estimated Costs of Human Genome Projects

Congress has primary responsibility for funding research through Federal agencies because of its responsibility for the national budget each year. Appropriating Federal funds for any special genome projects will therefore fall to Congress. These appropriations will express Congress' judgment regarding the relative value of genome projects. In setting appropriation levels, Congress will weigh the costs of the programs against their anticipated benefits (in economic and social terms) and will balance the value of proceeding against the costs of not doing so, as measured in lost benefits or opportunities.

Proposals for genome projects are intended to support research, but research needs are inherently unpredictable: Technological breakthroughs could dramatically diminish budget needs, and unanticipated obstacles could just as dramatically increase them. Estimates of near-term projects using existing technologies are necessarily more accurate than future projects that presume technological developments. Costs for some of the larger components, such as sequencing significant portions of human or nonhuman DNA, hinge on unit costs that are highly uncertain now and are rapidly changing due to technical advances (e.g., the cost of sequencing a single base pair of DNA). These uncertainties suggest that a 5-year budget plan is the best that can be produced, and projected costs for even the first 5 years might need to be substantially revised. The costs of human genome projects can be separated by components, although the boundaries between some of them are imprecise. The costs projected in this appendix are based on a process followed by OTA to generate estimates from internationally recognized experts.

OTA Cost Estimates

In order to better estimate potential costs of human genome projects, OTA held a workshop on August 7, 1987. At that workshop, there was apparent consensus on rough estimated costs of several components and confusion or disagreement about many others. A follow-up letter was sent to workshop participants and over 150 experts from executive agencies, universities, and corporations to confirm estimates made at the workshop and to expand them. Replies were received from over 70 persons. The revised cost estimates were externally reviewed by over 100 individuals and institutions in a draft report circulated in November 1987, and some minor revisions are based on comments received during this review. The resulting cost projections attempt to include most of the direct costs of research. They do not include indirect costs of university administration (although they do include administration in Federal agencies).

In some cases, it may prove possible to attract funding from the private sector—foundations, medical research institutes, or corporations. If so, Federal spending could be correspondingly reduced. In many cases, however, the Federal Government will eventually pay the full costs. If a company developed mapping and sequencing information or new instruments, for example, the first-and for a longtime the predominant—users would remain researchers funded to do biomedical research by the Federal Government. This would be the case for most technologies developed as part of human genome projects (use by researchers being the primary goal of the enterprise). A company's investment would thus be charged back to the government by charging for use of information or purchase of instruments by the research community. In some cases there may be a market for products outside the biomedical research community. If so, the private sector funds could indeed displace government funds. Funding from research foundations, medical institutes, and other philanthropies would also, as a rule, substitute directly for government costs.

There was strong consensus about the importance and feasibility of improving the research infrastructure (databases and repositories) and generating genetic and physical maps of human and nonhuman chromosomes; there was substantial uncertainty about sequencing strategies and their associated costs. It is agreed that the need for new technologies is paramount, but there is disagreement about how much it would cost to develop them or how such efforts should be organized.

Discussion in the following sections reviews costs by component.

Computers and Computational Methods

Cost estimates for the necessary persomel, research, and equipment are $12 million per year for the early years, increasing to 15 percent of the overall budget as it exceeds $80 million annually. This would be in

addition to continued support of existing databases and computer facilities. Spending should be relatively flat over time, because hardware will have to be purchased in early years and research will take an increasing proportion of the budget in later years. Hardware will have to be upgraded, however, so cost estimates are necessarily uncertain for future years. While it is logical to link computational needs to human genome projects, funding devoted to storage of genetic data and sophisticated analysis of DNA will prove important in molecular biology even if maps are not completed and other human genome projects are not funded.

Genetic Maps

Genetic mapping has been conducted for several years, and a rough map of human chromosomes already exists. Discussion at the OTA cost workshop centered on a map with two to four times the resolution of current maps. Subsequent letters and discussions have centered on a further increase in resolution, preferably such that a gene being studied would be separated from its closest DNA markers (on average) in only 1 of 100 family members. (Geneticists call this a l-entimorgan map.) Estimates based on existing procedures yield annual costs of $6 million per year for 5 years. Since this is an existing technology and there are already facilities to do the mapping, a startup period is not needed. Funds saved from new methods could be devoted to automating the processes so further refinements of maps in humans and other organisms would be easier to construct in the future.

The two principal groups constructing human genetic marker maps to date have not been federally supported. One has used private corporate funds, and the other has been funded by the Howard Hughes Medical Institute (HHMI). HHMI-sponsored work is a nearly direct substitute for goverment funding. Future work would be of greater magnitude, however, and may require Federal investment. In the case of work supported by Collaborative Research (the largest corporate group) and other companies, the Federal Government will probably pay for access to the probes either as a lump sum (to obtain access for all federally funded researchers) or indirectly (as federally funded researchers pay for access to individual DNA markers or mapping services).

Physical Maps

physical maps would be quite useful for future research. Ordered clone sets linked to them would be even more useful. Pilot projects on selected human chromosomes and on many lower organisms are in progress, and a useful set of ordered clones from all the human chromosomes may be feasible in the next 5 to 10 years.

Projections based on existing technolo~ yield costs of $60 million for a usefully complete set of ordered cosrnid clones over 5 years, New technologies may permit the creation of ordered sets of much larger DNA fragments (using yeast artificial chromosomes, YACS), and these would be extremely useful also. Costs of constructing ordered libraries composed of both cosmid and YAC clones are estimated at $7o million.

There are substantial uncertainties regarding both types of clones. Physical mapping of human chromosomes using cosmid clones has only begun in the last year, and therefore the rate and completeness of such mapping are highly uncertain. Mapping with yeast artificial chromosomes is much newer, although promising. The main uncertainty regarding YACS is not cost, but feasibility: If such mapping is possible, it would be substantially less expensive than mapping using cosmids (although cosmid maps might be needed for many research applications).

Ordered clone libraries are difficult to complete. Progress is rapid at first, but it is unlikely that a chromosomal region can be spanned without gaps between groups of continuous clones. Maps complete enough to be useful can be expected from several years' effort, but if truly complete maps are necessary, then efforts must be continued, perhaps at funding levels equal to those for initial construction. Half or more of the total effort maybe required for the last 10 percent of the maps. Clone libraries with gaps are quite useful, however, because a chromosomal region of interest is likely to be represented even in incomplete libraries.

Cost estimates start at $10 million for the first year (building on current Federal expenditures), rise to $2o million in $5 million increments over 2 years, and then drop to $10 million (with the proviso that continued higher funding may be necessary if complete maps are deemed essential).

Projects To Link Genetic and Physical Maps

Identifying the parts of DNA that carry the instructions for making protein and integrating them into genetic and physical maps would be very useful. Which stretches of DNA are actually used to produce protein varies with the tissue (many genes are expressed differently in different tissues or stages of development). The likely process would be to make DNA copies (cDNA) of the RNA that is translated into protein from a variety of tissues ('both healthy and diseased) and at

various stages of development. Locating cDNAs on physical and genetic maps would result in cDNA maps.

Such maps could be used to pick out protein-coding regions along stretches of DNA of unknown function. This would make the physical map much more useful, by highlighting regions of particular interest, and would provide a missing step in the search for genes whose approximate location had been determined by genetic linkage maps. DNA sequencing might also begin by using cDNAs to select regions likely to be of interest (because they are known to produce protein). Maps of cDNA would give clues to a gene's function if the pattern of expression related to a known biological process. Comparison of cDNAs from human and other organisms can give clues to function by relating expression to degree of evolutionary relatedness. If a genetic disease is located in a certain chromosomal region and cDNA maps show that one DNA segment from that region is transcribed only in the tissue affected by the genetic disease, then the gene corresponding to the cDNA is a good candidate for the gene causing the disease. Maps of cDNAs have been suggested by several groups [2,3,11,12,13,15].

The first step in constructing cDNA maps would be to collect and organize existing sets of cDNA clones. New sets of cDNA clones could be made from missing tissues, disease states, developmental stages, or organisms. The various cDNAs could then be located on genetic linkage maps and physical maps. The cost of this process is highly uncertain, in part because the number of genes in human and many other organisms is not known. Those specifically asked about this component estimated that its costs would likely range from $2 million to $5 million per year, depending on how much work could be done by merely cataloging existing cDNA clone sets; how many new sets would have to be constructed; how many organisms, tissues, developmental stages, and disease states would be used as sources; and the extent of genetic and physical maps. The costs of cDNA mapping would increase with the increasing detail of genetic and physical maps. OTA estimates start from a base of $2 million, increasing annually by $1 million increments.

Resource Material Repositories

Estimated costs of storing the clone sets linked to physical maps, cell lines for genetic research, and the various DNA analytical materials for genetic mapping originally ran to over $250 million. The largest component, dwarfing all others, was the cost of storing the DNA clones linked to physical maps. Such storage costs are virtually prohibitive, and these estimates were dropped, Subsequent discussions with experts on storage of materials for molecular biology, specifically with persons at the American Type Culture Collection, yielded storage estimates an order of magnitude lower. The estimates summarized in table B-1 are for collection and storage of clone sets. Costs of dissemination would be borne by users through user fees. Costs of collecting and storing mutant cell lines and DNA analytical materials (such as probes) have not been included.

Sequencing

There is little consensus on how much DNA sequencing should be done as part of genome projects, particularly whether a complete human reference sequence should be an objective. There is consensus, however, that sequencing technology is crucial and ripe for innovation. Cost projections should become easier in 2 to 3 years, as the first automated DNA sequencing machines are improved; massive sequencing would not begin in most schemes for several years,

Table B-1.—OTA Budget Projections for Genome Projects (millions of dollars, adjusted to 1988)

Component	Year 1	Year 2	Year 3	Year 4	Year 5
Computers and analysis	12	12	17	24	29
Genetic maps	6	6	6	6	6
Physical maps	10	15	20	20	10[a]
cDNA maps	2	3	4	5	6
Resource material repositories	1	2	3	4	5
Sequencing	—	—	15[b]	30[b]	45[b]
Quality control	—	1	2	3	4
Technology development	10[b]	20[b]	50[b]	75[b]	100[b]
Training	4	6	8	10	12
Aministration	2	3	6	9	11
Total	47	68	131	186	228

[a]May require upward adjustment for map closure.
[b]Subject to considerable uncertainty, depending on technical improvements, strategy, and unit costs.

SOURCE: Office of Technology Assessment, 1988.

except as part of pilot projects and for DNA regions known to be of special interest [6]. The debate about sequencing involves disagreement about the costs of sequencing per base pair, the amount of redundancy necessary to make a sequence useful, the expected pace of technological improvements, which laboratory preparation steps are included, and how much DNA would be sequenced as part of human genome projects (rather than through traditional funding mechanisms). Estimates of the cost of sequencing vary widely, ranging from several pennies to several dollars per DNA base [6,16], but there is some agreement that costs would drop to $0.20 to $0.30 per base pair by the end of 1988, based on existing technologies. Some of the discrepancy in the estimates comes from including different components. The costs of special cloning procedures, preparing DNA, use of reagents, technician time, and capital costs of instrumentation should all be included in cost estimates.

Judgments about which technologies to use and how much sequencing should be performed are best made each year by an advisory committee with access to technical experts. Such judgments would presumably be based on costs, the availability of material to sequence, and consensus on which regions to sequence first. For OTA projections, a few assumptions have been made. For the first 2 years' budget, sequencing would be covered as technology development—performed on lower organisms or human chromosomal regions of known interest —for possible sequencing on a larger scale. For years 3 to 5, it would be based on sequencing one small chromosome per year at $0.20 per base pair ($30 million per year, based on threefold redundancy and 50 megabases per year), permitting a phase-in period for implementation of the technologies. This estimate is for purposes of budgeting only, however, and could prove wildly high or low. If the technologies for cloning, preparing DNA for sequencing, and finally determining a DNA sequence become significantly cheaper, as some experts predicted at the OTA workshop, the *amount* of DNA sequenced at that cost could be increased. If costs remain high, only a limited amount of DNA could be sequenced, according to priorities set by the oversight board. After the fifth year, the budget could go up or down in proportion to need.

Quality Control and Reference Standards

The large amounts of map and sequence information and new materials created by human genome projects will be useful only if the information is accurate and resource materials are cataloged reliably. If there are many different groups involved in the efforts, problems of quality control could impede useful applications. The scope and magnitude of this problem will become clear only when the technologies are defined and the results of mapping and sequencing efforts begin to accumulate. Special budget allocations for comparing results from different groups or to establish measurement standards may become necessary. Budget needs for quality control will be nil in the first year and will grow in early years until they constitute 5 percent of the overall budget. For initial estimates, it is projected to grow by $1 million per year from a base of zero.

Technology Development

Investments in methods and instruments associated with genome projects are likely to lead quickly to commercial applications. The objective for technology development is open-ended, however, and it could be either the largest component or a relatively small fraction of genome projects. Responses to OTA letters and drafts showed no consensus on the proper budget. Many scientists familiar with industrial development encouraged higher figures, while academic molecular biologists set lower ones. A maximum figure of $500 million to be spent over 5 years was mentioned at the OTA workshop, in line with recommendations of a committee established by the Department of Enery (DOE). There was some support for the alternative of devoting 25 percent of the total budget to technology development [6]. Minimum estimates were for a steady state of $20 million to $30 million. Several individuals noted the importance of developing technologies early on, while recognizing the need to keep early budgets realistically low because a new research program would require the accumulation of trained personnel and pilot work to provide a foundation for later work.

The approach used in OTA estimates is to increase funding from $10 million the first year to a stable figure of $100 million by yearly increments. Funding for biological instrumentation centers under the National Science Foundation might account for part of this, and methods or instruments of great interest to industry might lead to some cost sharing with private firms. If so, Federal funding could be reduced accordingly. Technology development funding, like sequencing, is among the most flexible of the proposed projects and could be adjusted by the oversight board and the congressional appropriations process,

Training of Personnel

Training of investigators and scientific exchange among participants are crucial and would include grad-

uate and postgraduate fellowships, scientific workshops, and national scientific meetings. Some persons urge that fellowship funds be targeted to shortage areas, but others believe that targeted programs are less effective than untargeted ones for the best people in any relevant discipline. If training were targeted, it might include development of dual expertise in computers and molecular biology, organic chemistry and molecular biology, engineering and molecular biology, and clinical medicine and informatics or molecular biology. Training would also be needed for technicians, and for sabbaticals for scientists interested in shifting from their fields to genome projects. Workshops among participating groups and national symposiums to communicate results would permit rapid dissemination of new methods and insights. Exchange programs among industrial, national laboratory, and academic scientists would promote technology transfer. Training and personnel costs are estimated to merit 10 percent of each annual budget. For initial projections, funding might start at $4 million and increase yearly by $2 million.

Administrative Costs

Participants in the August 1987 OTA workshop estimated that 1 to 3 percent of each year's budget would be needed for administrative overhead. That estimate was subsequently increased to 5 percent in response to letters and after analyzing administrative costs at Federal research agencies. Administrative costs include operation of a national advisory board; oversight of databases, repositories, networks, and other services; setting instrumentation standards for cloning, mapping, and sequencing technologies; administration of grants and contracts; and other purposes. Some additional features would be unique to genome projects, for example, analysis of likely social impacts and ethical dilemmas created or intensified by genome projects. The need for such analysis has been explicitly noted in hearings and has been highlighted by research agency administrators and congressional staff. It could be obtained through grants to bioethicists, lawyers, economists, and social scientists for publications or workshops on various topics.

Summary

The costs of the components of human genome projects are projected in table B-1. These would start from a base of $47 million in fiscal year 1989 (if 1989 were the first year) and increase to $228 million in fiscal year 1993. It is not useful to project budgets beyond then, because technological development is so uncertain. The projected figures do not attempt to assign functions to particular agencies, merely to state overall direct research costs. Future budgets will need to be revised in light of actual appropriations.

History of Earlier Estimates

Perhaps the earliest evidence of a human genome project is found in a letter from Robert L. Sinsheimer, then Chancellor of the University of California, Santa Cruz (UCSC), to University of California President David Pierpont Gardner, on November 19, 1984. A potential benefactor had withdrawn support from a project, and Sinsheimer took the opportunity to provide a counterproposal that might interest the benefactor. In doing so, Sinsheimer suggested that a human genome institute be founded at UCSC, with startup costs of $25 million and an annual operating budget of $5 million. This was, in effect, the first cost estimate for a human genome project.

The letter from Sinsheimer to Gardner referred to an enclosure, later to be used as a basis for discussion at the May 1985 Santa Cruz Human Genome Workshop, in which the institute was formally proposed [71. The proposal assumes existing mapping technologies and a continued rate of development of DNA sequencing speed equal to the exponential increase of the past decade. The proposal then concludes that "within a few years, the human genome could be reduced to an ordered set of cloned fragments" and that "50 tech" nicians could approach completion of the [sequencing] project in 10-20 years." The proposal estimates the yearly support of each technician at $100,000, yielding an annual budget of $5 million and a total project cost of about $100 million. The proposal also calls for $25 million for startup facilities, and it distributes the operating money among the mapping and sequencing project itself (75 percent), developing techniques (10 percent), application to basic biology and medicine (10 percent), and education and training of students and other personnel (5 percent).

The Santa Cruz workshop displays similar optimism about the mapping aspect of a genome project, suggesting that a physical map could be completed by a 20-person group in 3 to 5 years [14]. The workshop also included discussion of a restriction fragment length polymorphism (RFLP), or genetic, map. This map could be achieved in "a few years" at a resolution finer than 50 centimorgans. Based on then current technologies, sequencing the 3 billion base pairs of the human genome was taken as "not feasible." The workshop went beyond the initial proposal and discussed details about the computer requirements for a project. There was, however, no explicit cost estimate for these details,

The next round of cost estimates came out of DOE's workshop in Sante Fe in March 1986. Appended to the workshop notes and the correspondence the workshop generated between the participants and DOE's Mark Bitensky was a cost estimate by Christian Burks of the Los Alamos National Laboratory. Burks calculates the person-years required for various aspects of the project, which, for a physical mapping and sequencing endeavor, including computer and administrative costs and assuming some sequencing advances, totals 3,505 person-years [5]. Allowing for hardware and overruns of his estimate, Burks concludes a genome project would cost between $0.5 and $2.5 billion.

The next major meeting, at Cold Spring Harbor, focused primarily on sequencing. The only estimate to issue from the discussion was the oft-quoted 30,000 person-years required to sequence the human genome one time through [8]. This estimate—translated into $3 billion by either $1 per base pair or $100,000 per person-year–was based solely on existing technology and was therefore obsolete within days of the conference, when the automated sequencer at the California Institute of Technology was announced [18].

By the middle of 1986, the Caltech sequenator had made it clear that advances in sequencing technology would drive costs down. HHMI's Informational Forum at the National Institutes of Health in Washington, D. C., continued to quote the 30,000 person-year estimate, but it also cautiously offered an estimate of 300 person-years, assuming a two-order-of-magnitude increase from automation [17]. The HHMI forum likewise gave a dual estimate for the physical map (200 person-years, or 30 to 40 with automation advances) [4], and for computer storage of sequence information ($0.30 per base pair, $0.03 with advances) [11.

Nine months later, DOE brought out its own cost estimates, presented as a yearly budget for a genome project. In the Health and Environmental Research Advisory Committee report, the subcommittee scientists estimate that sequencing, with redundancy for accuracy, would cost $60 million, assuming advances in automation [19]. Sufficient automation should be available 5 years hence [10]. The remainder of the budget is not described in detail, but it does specify that $500 million will go to various aspects of technological development, including mapping and informatics, assuming $100,000 per person-year of research [9]. The total for the DOE-proposed projects comes to $1.02 billion. Table B-2 presents a summary of estimates.

The National Research Council of the National Academy of Sciences established the Committee on Mapping and Sequencing the Human Genome, whose report was released in February 1988 [13]. That report represents the views of an exceptionally distinguished panel of experts from diverse scientific backgrounds, The panel members began their deliberations with widely differing knowledge of the state of gene mapping and divergent opinions about the merit of special research efforts. While writing the report, the panel reached a consensus that a special effort was merited and recommended additional funding of $200 million per year. This level would be reached over the initial 3 years. During the first few years, the budget would be roughly divided into $120 million for research in 10 or so multidisciplinary centers and numerous small research groups. Construction and materials would cost $55 million per year, and $25 million would Operate repositories, databases, training, and administrative functions. In later years, the budget would increase for dedicated production of map and sequence data. This $200 million annual budget would continue until at least the year 2000.

Appendix B Refemmces

1. Be]], George 1., comments at Informational Forum on the Human Genome, sponsored by Howard Hughes Medical Institute, July 23, 1986, at the National Institutes of Health, Bethesda, MD.
2. Berg, Paul, comments at Costs of Human **Genome Projects, OTA workshop, Aug.** 7, 1987.
3. Berg, P., Stanford University Medical Center, personal communication, January 1987.
4. Brenner, Sydney, comments at Informational Forum on the Human Genome, sponsored by Howard Hughes Medical Institute, July 23, 1986, at the National Institutes of Health, Bethesda, MD.
5. Burks, Christian, "The Cost of Sequencing the Complete Human Genome," Santa Fe Genome Sequencing Workshop, Mar. 3-4, 1986.
6. Costs of Human Genome Projects, OTA workshop, Aug. 7, 1987.
7. Edgar, Bob, Noller, Harry, and Ludwig, **Bob, "Human Genome Institute: A Position Paper, " enclosure in personal correspondence from Robert L. Sinsheimer, University of California at Santa Cruz, to David Pierpont Gardner (Nov. 19, 1984) distributed at the Santa Cruz Human Genome Workshop, May 24-25, 1985.**
8. **Gilbert, Waker, comments at Molecular Biology of Homo *sapiens*, Cold Spring Harbor Laboratories symposium, May 28-June 4, 1986.**
9. **Hood, L., California Institute of Technology, personal communication, June 1987.**
10. **Hood, L. California Institute of Technology, personal communication to David Guston, June 1987.**
11. **Karnei, R., and Stocker, J.T., Armed Forces Institute of Pathology, personal communication, December** 1987.
12. McKusick, V. A., and Ruddle, F.H., "Toward a Complete Map of the Human Genome," Genornjcs 1:103-106, 1987.
13. National Research Council, National Academy of Sciences, ***Report of the Committee on Mapping and Sequencing the Human* Genome (Washington, DC: National Academy Press, 1988).**

14. Notes and Conclusions, Santa Cruz Human Genome Workshop, sponsored by Department of Energy and Los Alamos National Laboratory, June 4, 1985.
15. Philipson, L., and Tooze, J., "The Human Genome Project," BioFutur, June 1987, pp. 94-101.
16. Roberts, L., "New Sequencers to Take on the Genome, " Science 238:271-273, 1987.
17. Smith, David, remarks at Informational Forum on the Human Genome, sponsored by Howard Hughes Medical Institute, July 23, 1986, at the National Institutes of Health, Bethesda, MD.
18. Smith, Lloyd M., Sanders, Jane Z., Kaiser, Robert J., et al., "Floumscence Detection in Automated DNA Sequence Analysis," *Nature, 321:674-678, 1986.*
19. Subcommittee on the Human Genome, Health and Environmental Research Advisory Committee, *Report on the Human* Genome Initiative, prepared for the Office of Health and Environmental Research, Office of Energy Research, Department of Energy (Germantown, MD: DOE, April 1987).

Table B-2.–Comparison of Genome Cost Estimates (millions of dollars (M) or person-years (py))

Source	RFLP map	Physical map	Sequence	Computing	Repository	Other	Total
UCSC[a] position paper 1 1/19/64		"a few years"	500-1,000 py				$25 M facilities
UCSC workshop 5124-5185	"a few years" (<50 cM)	60-100 py	"not feasible"				
Sante Fe workshop 3186		55 py	3,000 py	300 py + hardware		150 py administration	$500-2,500 M[a]
Cold Spring Harbor Symposium 5/28-6/2/66			30,000 py[b]				
HHMI/NIH Informational forum 7/23/86	200 py or 30-40 py[b]	30,000 py or 300 py[b]	$.30/bp or $.03/bp[b]				
DOE/HERAC 4/87			$60 M or 6,000 py[a]			$500 M technology	$1,020 M
NRC						$60 Mlyr 10 centers[a] $60 M&r grants and technology development for small groups $55 M/yr[c] facility construction (early years, decreasing later) $25 Mlyr administration, quality control, advisory committee functions	$3,000 M (over 15 yrs.)
OTA[d] 8/87-1/88	$30 M (1cM)	$70 M YAC and cosmid	$60 M not complete	$12 M min 15% of total	$15 M	$200 M/yr $10 M quality control $20 M linking projects 5% administration $255 M technology development $40 M training	$660 M (first 5 yrs. only)

[a]Assumes $100,000 per person-year.
[b]Not consensus figures but individual opinions.
[c]Money for facilities in early years would go to mapping and sequencing in later years.
[d]Estimates for first 5 years only. Does not assume complete reference sequence. For details, see text.
Abbreviations: DOE/HERAC—Health and Environmental Advisory Committee, Department of Energy; HHMI—Howard Hughes Medical Institute; NIH—National Institutes of Health; NRC—National Research Council; OTA—Office of Technology Assessment, U.S. Congress; UCSC—University of California at Santa Cruz.

SOURCES **UCSC:** personal communications from Robert Sinsheimer, Chancellor, UCSC, January 1987 and August 1987; Santa Cruz Human Genome Workshop (SCHGW), "Notes and Conclusions," June 4, 1985; and Bob Edgar, Harry Noller, and Bob Ludwig, "Human Genome Institute: A Position Paper," enclosure in personal correspondence from Robert L. Sinsheimer to David Pierpont Gardner (Nov. 19, 1984) and distributed for Santa Cruz Human Genome Workshop (May 24 to 25, 1985); **Sante Fe Workshop:** Christian Burks, "The Cost of Sequencing the Complete Human Genome," App. VI: Genome Sequencing Workshop, Sante Fe, NM, Mar. 3 and 4, 1986; **HERAC:** U.S. Department of Energy (DOE), *Report on the Human Genome Initiative,* Subcommittee on Human Genome of the Health and Environmental Research Advisory Committee, April 1987; **Cold Spring Harbor Symposium:** Walter Gilbert, comments at Cold Spring Harbor Laboratories Symposium, "Molecular Biology of *Homo sapiens,* May 28 to June 4, 1986; **HHMI/NIH symposium:** Remarks of George Bell, Sydney Brenner, and David Smith, at Howard Hughes Medical Institute Informational Forum on the Human Genome, July 23, 1986; **NRC:** National Research Council, National Academy of Science, Committee on Mapping and Sequencing the Human Genome, *Mapping and Sequencing the Human Genome* (Washington, DC: National Academy Press, Feb. 1988); **OTA:** "Costs of Human Genome Projects" Workshop, Aug. 7, 1987, with subsequent summary letter and review, October 1987, and external review of cost projections December 1987.