DOE/ER/61584-T1

The Human Genome Project: Information Management, Access, and Regulation

Technical Progress Report

For the Period
1 April - 31 August 1993

Submitted
10 September 1993

by

Joseph D. McInerney, Principal Investigator
and
Lynda B. Micikas, Project Director
BSCS
830 North Tejon, Suite 405
Colorado Springs, Colorado 80903
(719) 578-1136

Supported by Grant Number DE-FG03-93ER61584
from the
United States Department of Energy

List of Accomplishments During the Current Grant Period

Per the original application, project staff have accomplished the following tasks during the current grant period:

1. *Advisory committee meeting.* On 16, 17 April 1993, the project advisory committee convened at the Sheraton South Hotel, Colorado Springs, Colorado to set the conceptual framework for the project. The agenda and list of participants are attached as Appendices A and B; the summary of the meeting -- the conceptual framework for the project -- is attached as Appendix C.

2. *Review of the conceptual framework.* During May and June 1993, the conceptual framework was reviewed by the advisory committee and by the education committees of the American Society of Human Genetics, the National Society of Genetic Counselors, and the Council of Regional Networks of Genetic Services. It also was reviewed by two independent reviewers, Dr. Thomas H. Murray, director of the Case Western Reserve University School of Medicine Center for Biomedical Ethics and Dr. James M. Sikela of the Department of Pharmacology, University of Colorado Health Sciences Center at Denver, Colorado. A copy of the evaluation instrument is included as Appendix D; a summary of the reviews is attached as Appendix E.

3. *Creation of two prototype databases.* During June and July 1993, the project staff and representatives of Learning Systems Consultants, Inc. (LSCI) designed and developed two prototype databases for demonstration and use during the writing conference. The databases were developed on the basis of the general guidelines and activities proposed by the advisory committee and summarized in the conceptual framework for the project. The service contract with LSCI is attached as Appendix F; a brief description of the prototype databases is attached as Appendix G.

4. *Writing conference.* The writing team convened at BSCS headquarters from 19 July - 31 July to draft the experimental materials, using the conceptual framework as the blueprint. The agenda and list of writers are attached as Appendices H and I. A preliminary outline of the experimental materials is included as Appendix J.

5. *Revisions of the experimental materials.* The project staff continues to revise the products of the writing conference, with internal review and editing occurring with each iteration. The project time line calls for a full set of revised and edited experimental materials to be returned to the writers for their review and comment by 15 October, and for the final draft of the field-test materials to be ready for duplication by 15 November.

6. *Creation of a field-test version of the databases.* BSCS has contracted with LSCI to assist the project staff with the design and development of the field-test version of the

model databases. These databases will be developed on the basis of the prototype databases created under the BSCS/LSCI Service Contract dated 10 July 1993 and a set of general requirements and changes generated by BSCS staff following the writers' conference. A full conceptual description of the field-test version of the databases is attached as Appendix K.

The contract with LSCI is attached as Appendix L. The contract specifies deadlines of 15 October and 15 November for delivery of the pilot product and the field-test version, respectively. The contract also requires LSCI to work with BSCS to revise the field-test version of the databases on the basis of the feedback received during field testing. Delivery dates for final Macintosh and MS-DOS versions of the software are 1 May and 1 July 1994, respectively.

7. *Selection of field-test sites.* To date, nineteen teachers from across the United States have expressed interest in directing the five field-test sites available for the project. The list of potential field-test teachers is included as Appendix M. On 2 September 1993, BSCS mailed cover letters and background surveys (Appendix N) to each of the teachers listed; final selection of field-test sites is scheduled to be completed by 1 November. On 3-4 December, the field-test teachers and several writers will meet in Colorado Springs for the orientation session. A preliminary agenda for the field-test orientation is attached as Appendix O.

8. *Publicity.* The project has received publicity through a number of mechanisms, including:

   - announcements in science-education newsletters: *BSCS: The Natural Selection* and *News and Views*, published by the National Association of Biology Teachers;
   - announcements at BSCS Green Version and Blue Version summer institutes for high school biology teachers (July 1993);
   - a seminar (10 August 1993) by the principal investigator at the Social Science Education Consortium, Inc., Boulder, Colorado, for the History and Nature of Science and Technology project; and
   - a seminar (12 August 1993) by the principal investigator at a bioethics workshop held at Buena Vista, Colorado, and sponsored by the Colorado Association of Independent Schools.

The principal investigator and project director also have arranged to offer workshops on the project at the western area convention of the National Science Teachers Association, Denver, Colorado, 28-30 October 1993; at the annual meeting of the National Association of Biology Teachers, Boston, Massachusetts, 17-21 November 1993; and at the national convention of the National Science Teachers Association, Anaheim,

California, 30 March-2 April 1994.

BSCS plans to publicize the module on the INTERNET, through listserves associated with the National Association for Research in Science Teaching (NARST) and the Association for the Education of Teachers in Science (AETS), and others as appropriate. BSCS also is considering publicizing the project through K-12 NET, an international network of teachers who use telecommunications for the exchange of data and information.

Concerns About Publication and Distribution

During the final budget negotiations for this project, BSCS was required to eliminate all funds for printing and free nation-wide distribution of the print materials. The assumption at that time was that teachers would download those materials from the BSCS electronic bulletin board in the same manner as they would the program software, a procedure we also instituted to reduce the budget.

The principal investigator expressed his concern about these constraints at the time of the budget negotiations. This matter is extremely serious given that peer reviews of the proposal questioned whether reliance on the bulletin board would restrict use of the program to those teachers who are quite computer literate and who have access to phone lines and telecommunications in their schools. (The letter included as Appendix P summarizes these concerns.)

This concern has not abated as the project has progressed. Indeed, the advisory committee and the writers agree that the conceptual framework for the project, the structure of the database software, and the nature of the print materials combine to exacerbate these concerns. The writers especially are convinced that the materials they designed during the July writing conference will attain only limited use nation wide if the current plans for distribution of the software and print materials remain unchanged.

The project staff concurs with the writers and the advisory committee. We believe the materials as now designed are conceptually sound and will accomplish the goals of the project in fine style (especially after extensive field testing and revision), but they must first reach the nation's teachers. This latter assumption is in question if current plans prevail.

Without additional funds to ensure the proper distribution of the software and print materials, the DOE investment in the project to date may go for naught. This unsatisfactory prospect means that the project staff will pursue additional funds elsewhere; in addition, we request that the Department of Energy consider a supplemental grant of $85,000, or some portion thereof, to help ensure that its investment does not remain locked in the BSCS electronic bulletin board.

List of Appendices

Appendix A: Agenda for the first advisory committee meeting

Appendix B: List of participants for the first advisory committee meeting

Appendix C: Summary of the first advisory committee meeting (conceptual framework)

Appendix D: Evaluation instrument for the conceptual framework

Appendix E: Summary of the review of the conceptual framework

Appendix F: Service Contract with Learning Systems Consultants, Inc.

Appendix G: Description of the prototype databases

Appendix H: Agenda for the writers' conference

Appendix I: List of participants for the writers' conference

Appendix J: Preliminary outline of the experimental materials

Appendix K: Description of the field-test version of the databases

Appendix L: Contract with Learning Systems Consultants, Inc.

Appendix M: List of potential field-test teachers

Appendix N: Cover letter and background survey sent to potential field-test teachers

Appendix O: Preliminary agenda for the field-test orientation

Appendix P: Letter to Daniel W. Drell, Department of Energy, Washington, DC

APPENDIX A

Agenda for the First Advisory Committee Meeting

A G E N D A
of the
FIRST ADVISORY COMMITTEE MEETING
for
The Human Genome Project: Information Management, Access, and Regulation
at
Sheraton Hotel, Colorado Springs
2886 South Circle Drive
(719) 576-5900/Fax (719) 576-7695

FRIDAY, 16 APRIL -- Fremont Room

8:00 - 8:15 a.m. Welcome and introductions

8:15 - 8:30 a.m. Overview of project

- rationale
- intended audience
- product
- distribution
- project time line

8:30 - 9:45 a.m. Group discussion of the conceptual framework

9:45 - 10:00 a.m. Break

10:00 - 10:15 a.m. Group assignments and discussion of tasks (**Fremont and Avondale Rooms**)

- group 1 - science and technology: Gottesman, Hartung, Mural, and Murray
- group 2 - ethics and public policy: Bingman, Leonard, Reilly, and Rothstein

10:15 - 11:30 a.m.	Work on group assignments
11:30 - 1:00 p.m.	Lunch (**Goldcamp Room**)
1:00 - 3:00 p.m.	Continue group work
3:00 - 3:15 p.m.	Break
3:15 - 4:30 p.m.	Review of progress
5:30 p.m.	Dinner

SATURDAY, 17 APRIL -- Fremont and Avondale Rooms

8:00 - 10:00 a.m.	Conclude group work
10:00 - 10:15 a.m.	Break
10:15 - 11:30 a.m.	Group reports and discussion
11:30 - 1:00 p.m.	Lunch (**Foothill Room**)
1:00 - 2:00 p.m.	Conclude group reports
2:00 - 3:00 p.m.	General discussion

- project evaluation
- resources for writers
- miscellany

3:00 p.m.	Adjournment

APPENDIX B

List of Participants for the First Advisory Committee Meeting

PARTICIPANT LIST
of the
ADVISORY COMMITTEE MEETING
for
The Human Genome Project: Information Management, Access, and Regulation

Ken Bingman
Shawnee Mission West High School
8800 W. 85th Street
Shawnee Mission, KS 66212
school (913) 642-3900, ext. 162
fax (913) 831-1450

Kay Gottesman
Manager, User Services
Human Genome Data Base
2024 E. Monument St.
Baltimore, MD 21205
office (410) 955-0289
fax (410) 614-0434
E-MAIL
KSG2LIBRARY.WELCH.JHU.EDU

Walter Hartung
Nederland High School
County Road 1-30
Nederland, CO 80466
school (303) 258-3212

Claire O. Leonard, MD
Department of Pediatrics
University of Utah
School of Medicine
50 North Medical Drive
Salt Lake City, UT 84132
office (801) 581-8943

Richard J. Mural, PhD
Senior Staff Member
Biology Division
Oak Ridge National Laboratory
P.O. Box 2009
Oak Ridge, TN 37831-8077
office (615) 576-2938
fax (615) 574-1274

Jeffrey C. Murray, MD
Division of Medical Genetics
Department of Pediatrics
University of Iowa Hospitals & Clinics
Iowa City, IA 52242
office (319) 356-3508
fax (319) 356-3347

Philip Reilly, MD, JD
Shriver Center
200 Trapelo Rd.
Waltham, MA 02254
office (617) 642-0230

Mark Rothstein, JD
Law Foundation Professor of Law
Health Law and Policy Institute
University of Houston Law Center
Houston, TX 77204-6381
office (713) 743-2101

BSCS Staff (719) 578-1136

Joe McInerney, principal investigator
Lynda Micikas, project director
Phil Goulding, technical specialist
Wilbur Bergquist, evaluator
Pam Van Scotter, editor
Kathy Winternitz
Dee Miller, secretary

APPENDIX C

Summary of the First Advisory Committee Meeting
(Conceptual Framework)

The Human Genome Project: Information Management, Access, and Regulation

SUMMARY
Advisory Committee Meeting
16, 17 April 1993

Biological Sciences Curriculum Study (BSCS)
The Colorado College
830 North Tejon, Suite 405
Colorado Springs, Colorado 80903

Supported by Grant Number DE-FG03-93ER61584/A000
from the
United States Department of Energy

The Human Genome Project:
Information Management, Access, and Regulation

Summary
Advisory Committee Meeting
BSCS, Colorado Springs, Colorado

INTRODUCTION

On 16 and 17 April 1993, the advisory committee met to outline the conceptual framework for the project. The specific objectives for the meeting were to:

a) identify the major concepts in biology, database construction, and ethics and public policy that should pervade the module;

b) outline the content for the teacher's narrative;

c) outline five days of classroom instruction that address the scientific, technological, ethical, and public-policy aspects of electronic genome databases;

d) identify database software that will meet the needs of the project; and

e) specify resources that will be helpful to the writers and the project staff.

After an introduction and overview of the project, the committee turned its attention to the foregoing tasks. The outcome of the committee's work is summarized in the pages that follow.

I. MAJOR CONCEPTS

A. Concepts Related to *Biology* and the HGP

1. There is a relationship between genotype (DNA) and phenotype, but the relationship is not always linear and direct. We must beware of the pitfalls of reductionism and determinism.

2. There is actually no such thing as "THE" human genome, except at the level of map position (i.e., genes for specific traits can be identified at specific loci). Although we can describe a "generalized" human genome, variations in specific sequences (i.e., allelic differences) allow DNA from each of us to be identified as unique.

3. The Human Genome Project will help us chronicle intra- as well as interspecific variation. Data from the HGP will reinforce our understanding of the basic biological principles of unity and diversity.

4. Genomic variation provides an historical record of evolutionary change and relatedness through descent.

5. Conserved sequences provide evidence of evolutionary relatedness and possible functional importance.

6. Most genetic change is neutral. Most genetic variation does not cause disease.

7. The vast majority of the genome is non-coding.

8. A knowledge of gene structure allows us to search for genes in DNA sequences.

9. Mapping and sequencing are not the end of the research story. Although the HGP will provide the data required to identify most human genes, it will not, in itself, explain how these genes interact with each other and with the internal and external environments to generate structure or function.

10. It is not enough to sequence only the human genome. The HGP also is concerned with sequencing the genomes of several other well-studied organisms (e.g., bacterial, nematode, fruit fly, mouse, and *Arabidopsis* genomes). Comparison of these genomes to that of humans will help scientists find human genes more easily, and help deepen our understanding of how genes function.

11. The HGP is an international effort, but only a few countries are heavily involved. This raises interesting questions about sharing data with those nations that are not involved.

12. The HGP will generate map and sequence data at a rate and volume that is of a different order of magnitude (i.e., of a different scale) than we have experienced to date. This information must be collected, stored, analyzed, and made accessible to other researchers if it is to be fully useful.

B. Concepts Related To *Databases* and the HGP

1. The enormous quantity of information expected from the HGP will require that we develop new approaches to information collection, storage, and processing. Those working on the HGP might be able to benefit from the work of information scientists in other disciplines such as meteorology or space-based planetary imaging.

2. Genomic information is currently being stored in two basic types of databases. Research databases organize and store aggregated information about the

"generalized" human genome (e.g., catalogs of genes, details of the human map, lists of DNA and protein sequences). Registry databases, on the other hand, store personal genome data that might be used to record genetic conditions or susceptibilities, or to identify individuals based on their specific genetic profiles.

3. Electronic genome databases are powerful resources that can help us answer biological questions, especially about similarities and differences.

4. The content and structure of a database determine the kinds of questions one can ask of it. Individuals concerned with designing and building databases to store genomic data must consider carefully what data to include and how those data can best be organized so researchers can ask appropriate questions. Databases are periodically redesigned so new questions can be answered.

5. The usefulness of a database also is limited by the questions the researcher asks of it.

6. An early use of electronic genomic databases will be to assist finding genes related to human health and disease.

7. Finding a gene is only part of the biological story. Once a new gene is found, scientists still need to identify its function and trace its evolutionary history.

8. Databases can have errors.

9. Databases cannot be protected completely. The more widely accessible genomic data are made (and, therefore, the more useful they become), the less they can be protected.

10. The information stored in databases can be misused in a variety of ways. Much

of the possible misuse is unrelated to health and medicine.

11. Our increasing knowledge of the human genome brings with it increased personal and collective responsibility.

C. Concepts Related To *Ethics and Public Policy* and HGP Databases

1. The HGP raises a number of important questions that will spur research on social and ethical questions. These questions, which will require ethical analysis and policy determination, include:

- issues of privacy and confidentiality
- issues of accessibility
- issues of autonomy and paternalism
- questions about what it means to be "normal"
- questions about justice and both fair and unfair discrimination
- questions about the interface of personal and public good

2. These issues are not unique to the HGP, but the HGP may make us more aware of them, and, because of their magnitude, may force us to deal with them sooner than we might have otherwise.

3. A major question facing policy-makers will be who should have access to genetic data. The answers to this question may differ for research databases as opposed to registries.

4. A related question may be how directly available to the public these technologies should be. Both questions recognize the potential problems associated with the public's increasing ability to acquire health-related information without going through the health-care system.

5. The rapid development of commercial applications of knowledge gained through the HGP may make ethical and public-policy questions particularly pressing.

6. The HGP raises a number of ethical and public-policy issues that force us to confront many of the tensions that exist inherently between institutions (e.g., the state and the family) and among individuals. Such tensions include:

- conflicts between the public good and the individual good
- tensions between the state's right to know and the individual's right to privacy
- tensions between an individual's duty to act and the same individual's duty not to act
- conflict of interest among individuals in a family (genetics involves families)

7. Public policy is sometimes established by law (*de jure* public policy). In the absence of specific laws regarding the HGP, individuals, and sometimes institutions, may act as they choose (*de facto* public policy). Institutions often operate within regulatory constraints that prevent them from acting as they choose.

8. The public should be involved in helping to make public policy. To fail to act is often to make a decision.

9. The public's ability to participate in discussions of these issues requires education at a variety of levels, including precollege education, education of the adult public, education of medical and allied-health personnel, and education of policymakers.

II. CURRICULAR AND CLASSROOM CONSIDERATIONS

1. Key objectives of the project are a) to engage students and teachers in an analysis of the scientific, technological, ethical, and public-policy aspects of electronic genome databases; b) to highlight major principles in biology, the nature and methods of science, the nature and methods of technology (especially electronic databases), ethics, and public policy; and c) to provide a sound, workable classroom model for analysis of ethics and public policy.

2. The audience for the teaching materials is students and teachers of introductory high school biology. The vehicle for instruction will be software and print materials (approximately 100 pages).

3. The new module assumes that students have covered the basic genetics portion of the typical, first-year biology program, especially basic transmission genetics, the basic structure of DNA, and the basics of the central dogma of protein synthesis.

4. The new module must be designed as a "stand alone" unit of instruction that neither assumes nor requires completion of the earlier module (*Mapping and Sequencing the Human Genome: Science, Ethics, and Public Policy*). We should consider using portions of the first module as necessary, especially those portions that address the basics of ethics and ethical analysis.

5. It will be important to make it clear to teachers that the new module is not intended as a replacement for the earlier module, but as an extension of these issues in new directions.

6. Like the earlier module, the new materials should be designed to be accomplished over five consecutive days (class periods) of instruction (plus homework assignments). A possible allocation of time might devote two days to questions of database design and

access, and three days to a consideration of the ethical and public-policy questions such databases raise.

7. The core activities of the new module should be designed for primary use in the biology classroom. However, the module also should provide suggestions for extension activities that would allow more detailed study of other issues in biology (e.g., use of the information in electronic databases to trace evolutionary relationships) or related study in other disciplines (e.g., in social studies, literature, and computer science). Recognizing that science teachers will have to take the initiative if this type of collaboration is to occur, the writers might consider preparing detailed suggestions for how biology faculty can interact with faculty in other disciplines on these issues.

8. Teaching and learning should focus on inquiry, with the students doing the work through hands-on activities such as manipulation of model or actual databases, and small-group or classroom discussion.

9. The instructional activities should require students to interact repeatedly with the specially designed databases that will accompany the module.

10. Computer-based exercises should be written assuming a student/computer ratio of approximately 4:1. The instructional materials should discourage teachers from conducting these exercises as demonstrations, and should include suggestions about how to maintain student involvement and time-on-task in classrooms where only one computer is available.

11. The instructional activities should use examples from a variety of racial and ethnic groups to help ensure that all students see the issues raised by the HGP as relevant.

III. OUTLINE OF THE TEACHER NARRATIVE (25 pages)

Section 1: Scientific and Technological Aspects of the HGP

Part 1: What is the HGP?

A. What is the HGP and how was it conceived? How is it funded and organized? (Perhaps revise material from the current module.)

B. Who works on the HGP and how is the work accomplished? (PhDs, lab techs, animal-care people; vignettes; images of who is doing the work and how; what is manual and what is automated; the time line of the technology; robotics)

C. What do we hope to learn from the HGP? The HGP will map and sequence all of the estimated 100,000 human genes, as well as the genes from several other organisms. The information generated by the project will

- have applications in health and medicine

- lead to new insights into basic biology (e.g., function of non-coding sections of the genome, gene regulation, evolution, development)

- spur the development of new technologies (e.g., sequencing technologies, techniques for information storage)

D. What have we already learned from the HGP? (brief summary of interesting/surprising/valuable discoveries already made as a result of

sequencing activities)

Part 2: What kinds of information does the HGP generate and how is this information stored?

A. The HGP generates information about the location of coding and non-coding regions on chromosomes and the sequence of bases in these regions. Already (by 1994) X amount of information has been collected. How much is expected?

B. What is an electronic database and how are such databases used? Why is electronic storage the method of choice?

C. How are research databases different from registry databases? (examples of research databases such as GDB, PIR, GenBank, OMIM/MEDLINE and registry databases used in hospitals, for forensic purposes, or by the military)

D. What are some specific databases associated with the HGP and what kinds of information do they contain?

[table to summarize characteristics of databases: who uses them and how, location, funding, who edits and updates them, structures, quality assurance, read/write issues, protection, size, complexity, speed of searching]

[blackline master for teachers to show how the CF gene is represented in various databases]

Part 3: How Are These Data Used?

A. How do researchers find genes?

- overview of gene structure (coding regions, regulatory sequences, simple versus complex genomes, translational punctuation, mutations in introns versus mutations in exons)

- how scientists locate genes [techniques for mapping genes from first module; mapping at the gene level from first module; importance of looking at similarities among genes (e.g., linkage groups among similar species, gene families, conserved sequences indicating relatedness and functional importance)]

B. How do clinicians use these data?

- diagnosis of suspected genetic problems; genetic counseling; screening and testing

- clinicians and basic researchers exchange interrelated information

C. How are registry databases used?

- use of registry databases in hospitals, by the government, insurance companies, the military

D. What problems are encountered in database use?

- problems of design

- problems of data accuracy

- problems of privacy

- problems of misuse

Part 4: What kinds of biological questions can databases help answer?

A. Where are genes located on chromosomes?

- map databases (genetic linkage maps and physical maps; the complete sequence as the ultimate physical map)

- location of human genes as compared to similar genes in other organisms

B. How can we recognize a gene?

- open reading frames

- conserved sequences

C. How are genes organized?

- location of coding and non-coding regions

- specific domains related to specific protein structure

D. How are genes related to human disease?

- the extent of variation in sequences

- variations that do not cause disease (functional changes)

- DNA markers and variations associated with disease

- candidate genes

- variations in specific human populations

- relationship between human genetic/evolutionary history, current environment, and multifactorial disease

E. What do similarities and differences in sequences tell us about evolution?

- linkage groups in different species

- conserved sequences (functional importance)

- gene families (cytochromes, globins, T-cell receptors)

- sequence comparisons, mutations rates, and phylogeny (mitochondrial DNA)

Transition to Section 2:

Biological knowledge derived from the HGP and databases has consequences for ethics and public policy.

Section 2: Ethical and Public Policy Aspects of the HGP and Related Databases

Part 1: What kinds of ethical and public policy questions will information from the HGP raise? (Revise from earlier module; focus particularly on information storage and access.)

A. Questions related to organization and access

B. Questions related to medical and economic issues

Part 2: What specific ethical principles are at issue in these questions?

A. Privacy (the right to keep certain information from being disclosed)

B. Confidentiality (nondisclosure by others to people who shouldn't know)

C. Autonomy (the right of individuals to control decisions that affect them directly)

D. Paternalism (the ability of others to make decisions affecting an individual)

E. Justice (like situations treated alike; dissimilar situations treated differently)

F. Discrimination (treatment that is different based on actual or perceived differences)

- fair discrimination

- unfair discrimination

Part 3: How do we subject these questions to systematic ethical analysis?

A. The features of ethics (revised from earlier module)

B. The role of argument in ethical inquiry (revised from earlier module)

Part 4: What legitimate public policy outcomes (goals) are at issue in these questions?

A. Justice

B. Economic efficiency

C. Equal access

D. Pluralism (cultural diversity)

Part 5: How are public policy decisions made?

A. Public sector

B. Private sector

C. Individual

[Table showing how these ethical/public policy principles are addressed in the student activities.]

	Felon Databank	Loan Application	"Smart" Card
Privacy	x	x	x*
Confidentiality	x	x	x*
Autonomy/Paternalization	x*	x	x
Justice/Discrimination	x	x*	x
Decision Process	x	x	x*
Efficiency	x*	x	x
Equal Access			x
Respect for Pluralism	x		x

*Each exercise involves several issues; starred points are emphasized by discussion questions.

IV. CLASSROOM INSTRUCTION (75 pages)

Overview:

Classroom instruction will cover five consecutive class periods. All activities will require students to interact with the databases we design for the program, first, the "National Genome Database" (NGD), and second, the "Local Genome Database" (LGD).

We will organize the NGD as a set of GenBank or GenBank-like entries, including, for each gene, such information as a 25-100 base consensus sequence, a summary of function, and other appropriate information (e.g., a list of references, the amino acid sequence, other). We will organize the data by species, and will include data on approximately ten genes, as they appear not only in humans, but in other selected organisms as well.

The LGD will contain data on the fictitious population. Each record will correspond to one person and will include a "name/number" field for identification, fields for the base sequences of the ten genes, and a field to indicate gender.

Because we will make the database available on our BBS, we need to be cautious about the length of the file that the teachers will download. If it is too long, the teachers may not have time to pick up the file, even though there is no cost to them. We will use self-extracting compressed files whenever possible. We also must be aware of the length of the programs that the teachers will use. We will use at least two specialized programs that will allow teachers and students to compare sequences.

The activities are organized as follows:

Days 1 and 2: Futuristic "2015" scenario, in which students reenact a supposed "DNA analysis" of the members of a small town. This activity gives students opportunities to manipulate a model genomic database, search for polymorphisms, suggest family relationships on the basis of genomic data, formulate and test

hypotheses about anomalies, and consider the possible implications of attaching names to sequence data in a research database, effectively converting a research database into a registry.

Days 3 through 5: Three individual activities related to ethical and public policy aspects of electronic databases. On Day 3, students will examine questions about the use of registry databases to store information about specific individuals. On Day 4, students will consider the use of genetic information for purposes unrelated to health care. On Day 5, students will discuss the use of medical data cards. These activities will give students opportunities to consider ethical and public policy questions related to privacy, confidentiality, autonomy and paternalism, justice and discrimination.

Days 1-2: DNA Analysis in the Year 2015

This activity is placed in the year 2015, in a small town with an eager 10th grade biology class (played, of course, by the current students). The fictitious students are studying genetic variability, and decide, because they have inexpensive DNA sequencing tools available in the lab, to collect actual samples of hair from the townspeople for analysis.

Most of the students collect samples from family members and friends, always with the permission of the donor. One student, however, who works at the hair salon, collects samples from the floor (without permission). Several days and some 30 strands of hair (plus follicles) later, the students have the information required to build a small research database (the LGD) containing the nucleotide sequences of 10 specific genes from each of 30 individuals. (The background material provided to teachers on this activity will address the problem of diploidy as it relates to this type of analysis.) Although two students in the class keep a master list that matches names to sample numbers, data entered into the database are identified only by number.

Exercise 1: Students might be given the opportunity to do some data entry by entering the nucleotide sequences for several of the genes for one of the 30 individuals. Using the search software provided, students could compare their data to that already loaded in the database for that individual. This would give students a concrete sense of the structure of the database, and allow them to consider questions related to the accuracy of the information stored in a database.

At this point, the fictitious biology teacher (unaware that some of the samples have been collected without permission) asks his or her students to consider how their data could be used to study genetic variation and relatedness in the town. The students discuss the problem and decide that they will analyze one gene in detail.

Exercise 2: Students are asked how they would decide which of the ten genes likely would be most useful if the class wanted to ask questions about genetic similarities and

differences, or if the data were going to be used to identify individuals. Students are helped to understand the importance of variability and are introduced to the concept of polymorphism. They are asked to examine their data for the presence of polymorphism and, acting as the fictitious class, to come to agreement about which gene they should analyze in detail.

Once the students have decided which gene to analyze, they compare the nucleotide sequences in their database for intraspecific variation, and, from their findings, construct a set of probable "family" clusters of sequences. This allows them to create a genetic picture, not only of individual family units, but also of the town (for example, they see that two families in the town are related).

Exercise 3 (Homework): The class is divided into small groups, and each student is given a print copy of one person's nucleotide sequence for the gene in question. (Samples are distributed so that each group represents one family unit.) Students are told that the next day they will be asked to compare their sequences with those of others in their group to determine probable family relationships. Students are asked to prepare for this by taking print copies of their sequences home and analyzing them for the presence of particular nucleotide patterns (e.g., locate and circle all A-T sequences, locate and circle all A-T-C sequences, locate and circle all A-T-C-A sequences, etc., up to 8 nucleotides). Students are asked to keep track of the time they spend completing each search.

Exercise 4: On Day 2, students discuss the difficulties of determining sequence similarities when large numbers of nucleotides are involved. Students repeat their searches on the computer, and extend them to determine relationships within families and within the town. Students discover two sequences that do not seem to fit the genetic picture they have developed and are asked to generate hypotheses about the possible identity of these two "outliers."

Exercise 5: Students test their hypotheses about the anomalous sequences using information in the NGD. They discover that one of the two sequences is human, and recognize that the lack of proper "fit" with a specific family suggests that the child was adopted. They discover that the second anomalous sequence is not human, and, looking across the species' sequences available, discover that it comes from a dog. Students discuss some of the possible sources of error in their database (e.g., entry errors, sampling errors, fraud).

By now, word of the fictitious students' work has reached the town council, and the students are asked to consider attaching names to the sequences in their database, effectively changing their "research" database into a registry. Town officials point out a number of advantages of such an action, but acknowledge that it might also trigger significant controversy. The students debate the issue, and, finally, reach a decision.

Exercise 6: Students consider the advantages (e.g., possible identification of the perpetrator of a future crime, identification of children in the event of kidnapping) and the disadvantages (e.g., will the family with the "odd" person agree, possible violation of the privacy of people who gave consent to have their DNA sequenced without fully understanding the implications, possible violation of the privacy of people who did not give consent, possibility that the gene they have been analyzing is or will be found to be related to susceptibility to a disease) of converting their research database into a registry. The activity ends with a vote (each student must make a choice), which stands as the decision made by the fictitious class.

Student Objectives:

After completing this activity, students will understand

- that although we can describe a "generalized" human genome, variations in specific sequences allow DNA from each of us to be identified as unique.

- that the HGP will help us chronicle intra- as well as interspecific variation.

- that not all genetic change has negative consequences.

- that the HGP will generate massive amounts of information that must be collected, stored, analyzed, and made accessible for it to be fully useful.

- how research databases differ from registries.

- that electronic genome databases are powerful tools that can help us answer biological questions, especially about similarities and differences.

- that databases can have errors.

- that our increasing knowledge of the human genome brings with it increased personal and collective responsibility.

- that the HGP raises a number of important issues that will require ethical analysis and policy determination.

Day 3: Felon Data Bank

A few weeks after the biology class completes its database, someone breaks into the local jewelry store and gets away with $X in precious stones and gold. The police find no clues, except for a broken window with some dried blood on it. The police send the pieces of glass to the forensic lab for analysis (especially DNA profiling).

Because felons are court ordered to give blood for DNA profiling, there presently exists a felon data bank. The bank may include a DNA fingerprint, a profile, or frozen white blood cells for future analyses. The DNA can be analyzed for thousands of different base sequences. The police report that they are going to access the felon data bank to try to identify the person involved in the robbery.

Exercise 1: The class is divided into small groups, which discuss the following questions:

- Is it right to require felons to give blood? Are their rights of privacy being violated? Should all felons in prison and out be DNA profiled? How about individuals who have been arrested, but not convicted?

- Who should have access to such a data bank? How should the data be used? When a crime is committed and DNA is collected, should the Felon Data Bank (FDB) be accessed to identify the criminal? Should the FDB be checked when one is applying for a driver's license? for jobs, especially where employees handle money? loans? scholarships? security clearances? passports? voting? before one can run for political office?

When a search of the felon data bank fails to turn up a match, the police turn their attention to the people in the town, and the chief asks the members of the biology class to run the DNA profile against their database to see if they can find a match. (Note: If the class decides to convert its research database into a registry database, Exercise 2 can proceed

immediately. If the class decides not to create a registry, then the students should be asked whether the existence of an actual unsolved crime is sufficient justification for changing their vote. If they still say "no," Exercise 2 can nonetheless be completed, with the students acting on behalf of a fictitious class that did decide to create a registry.)

Exercise 2: The students return to their small groups to compare the profile they are given to those in the database. To their surprise, they find that the profile matches that of a sample taken without permission from the floor of the barber shop. Students discuss 1) the significance of what they have found (e.g., does it establish guilt? how sure are they of the identity of the person from whom the sample was taken? how sure are they that their data entry was accurate? even if the students are sure that the police would take new samples from the person they identified before making any accusation, are they willing, on the basis of the evidence they have, to put a possibly innocent person through that public embarrassment and anxiety? does it matter that the original sample was taken without permission?) and 2) what they should do with their information (e.g., do they give it to the police?), as well as 3) what they should do with their database (e.g., is it, perhaps, too dangerous to keep? what kinds of non-science related use might be made of it, and can the students assure the rest of the donors that the information it contains will remain confidential and not be misused?)

Exercise 3: Students are asked to consider one or more of the following additional questions: Should the felon data bank be converted to a research database so all the felon DNA can be analyzed for a gene or DNA sequence that predisposes a person to criminal activity? If we find a "criminal sequence," should everyone with a criminal record be profiled to validate the finding? Suppose the military were to test incoming recruits for the presence of this sequence, and found that 10 percent of everyone tested carried the sequence, although only 1 percent of this same group were ever convicted of a criminal offense. How would you explain this finding? What questions might such a result raise about the use of this kind of information? Should prenatal testing for such a gene be allowed? If prenatal testing is positive, should the pregnancy be terminated? Could

felons use the presence of such a gene as a defense in court? Would you date/marry someone if you found out he or she had a "criminal" gene? A "violence" gene? What other questions does such profiling raise about privacy, confidentiality, justice, and assumptions about the biological basis of behavior?

Student Objectives:

After completing this activity, students will understand that

- the HGP raises issues of privacy, confidentiality, and accessibility.

- answers to questions about who should have access to genomic databases may differ for research databases as opposed to registries.

- there are inherent tensions between the state's right to know and the individual's right to privacy.

Day 4: Loan Application

A student in the fictitious "year 2015" town is almost 16, an A student, and has been working as a checker-bagger for 2 years. The local car dealer has a used car that the student hopes to buy for $3,500. His dad says he will cosign for a loan.

One of the questions on the loan application is: "Have you or your cosigner ever been denied life insurance for health reasons?" The father answers "yes," and the loan is denied because the lending agency has a policy of automatic refusal for anyone who answers this question in the affirmative.

The student and his father decide to protest. The father contacts the Medical Information Bureau (MIB) to find out what in his file may have led to the earlier insurance denial. MIB records show that this father's mother (the student's grandmother) died of Huntington disease. His father believes, however, that the grandmother died of Alzheimers, not Huntington.

Exercise 1: Students are asked to decide whether the lender's refusal was fair, and whether this represents a proper use of medical information. Can information in the insurance databank be corrected? Who has access to health care records? Is access sold? Does information about you belong to you or to someone else? What assures that the information in such databases is accurate?

Exercise 2: Students are asked to access the clinical information in the software (or to check their textbooks) to discover why the difference between Huntington and Alzheimers would be so important. The students are asked to write a paragraph to describe what the fictitious student might do to get his car, and, specifically, how the HGP might help.

The student decides that genetic testing for HD is needed to correct the MIB data, but to have the test accomplished by a certified laboratory would cost $500. His dad's hair sample is still in storage in the biology laboratory, and he suspects that the class could sequence it for him

if he asked.

Exercise 3: Students consider whether it is likely that the results of their analysis could be used. The student's father also had blood drawn for DNA analysis when he entered the army. Can it be used? Because it is expensive to have this test completed by health professionals, should over-the-counter tests (like pregnancy tests) be developed for common conditions? If so, how might people react if they find out that they have or may develop a serious disease? Should the developer of such tests have any responsibility for education and counseling?

At the father's request, the biology students sequence his DNA and compare the sequence to the known HD sequence. At the same time, they sequence the gene from all of the other samples they have.

Exercise 4: Groups of students compare the new sequences to the sequence of the HD gene stored in the NGD. They discover that the father does have the HD sequence. They also discover that the father's brother (the student's uncle, someone whose hair was taken without permission) carries the sequence as well. By chance, the student's hair was not one of the samples originally collected, so without direct testing, they cannot determine whether he also carries the HD gene. The students are asked whether, if they were the son, they would choose to be tested also. What implications might knowledge about his condition carry (e.g., for insurability, for employment, for marriage)? This section can describe the current procedures for HD testing.

The student's father is an accountant; the student's uncle is a bus driver. The students are asked to decide whom they will tell what. Will they tell the father that he carries the HD gene? Will they approach the uncle (who doesn't even know that the students have a sample of his hair)? Does the uncle's employer have a right to know that he carries the HD gene? If he learns that he has the HD gene, does he have an obligation to inform his employer? Do the students have an obligation to let the employer know, regardless

of whether the uncle is informed? Should the fictitious class have been permitted to do this analysis, given its potential problems? What if the sequence in question revealed one allele for sickle cell disease?

Student Objectives:

After completing this activity, students should understand that

- the HGP raises issues of autonomy and paternalism.

- there are a number of potential problems associated with the public's increasing ability to acquire health-related information without going through the health-care system.

- there are sometimes tensions between an individual's duty to act and the same individual's duty not to act.

Day 5: "Smart" Card

Some city health departments currently require that patients using public health clinics carry optical reading cards that encode their identification and their medical history (e.g., allergies, immunization, blood type, current medications or conditions). Similar cards are widely used in France. Such cards are used as a fast source of health care information about an individual, especially when the person is unable to provide it himself (e.g., if the person is unconscious, or cannot speak English).

The town council of the "year 2015" town, enamored of the power of the genetic information the students can derive from their hair samples, proposes to require that anyone who will be employed by the town, or anyone who requests treatment from the public hospital, have a smart card containing not only a full medical history, but a full genetic profile as well. For example, a person's card might show that he has genes that make him susceptible to various medical conditions (e.g., Marfans/heart disease/hypertrophic cardiomyopathy/alcohol susceptibility/mental illness/ADHD). The school board hears about the proposal and suggests that other kinds of data be encoded on the cards as well (e.g., IQ quartile, SAT scores), and that all school-age children be required to carry them.

Public hearings on the proposals are scheduled, but public understanding of the issues involved and interest in participating in the decision-making process seems very low. The members of the 10th grade biology class decide that they cannot sit back and allow the townspeople to "make a decision" by failing to act. The students ask the council for time to speak, and demand that their parents attend the council meeting to hear the debate.

Exercise 1: Students break into groups and prepare short statements to deliver to the town council. These statements might address such questions as 1) where would these data come from and how would they be collected (e.g., Where do these data come from in the places that currently use them? Is all screening voluntary? How about the newborn screening that occurs in hospitals? Should people have to give informed consent for the screening to occur? for the encoding to take place?); 2) Who should have access

to these data and why (e.g., Should such data be considered in hiring decisions? in decisions about insurability? in decisions about education and about "tracking" children in schools?); 3) Who should decide what is to be on the card? (e.g., Does all health information have to be on it? Who should decide what genes to include? Does the use of these data for non-medical purposes create the incentive for people to lie?); 4) whether such a program should be mandatory; and 5) whether information about mental health should be treated in the same manner as information about physical health; and 6) the extent to which such data, especially genetic data, define the person, and the danger that the public's lack of understanding coupled with the easy availability of such data might encourage reductionistic and deterministic attitudes about human life.

Student Objectives:

After completing this activity, students should understand that

- although there is a relationship between genotype and phenotype, the relationship often is not linear and direct.

- the information in databases can be misused in a variety of ways

- these issues are not unique to the HGP, but the HGP may help us be more aware of them.

- public policy is sometimes established by law, and sometimes by public action in the absence of law.

- the public should be involved in helping to make public policy.

- the public's ability to participate in a discussion of these issues requires education at a variety of levels.

As with the activities in the first module, the student discussions of ethics and policy will be constructed to fulfill the requirements of sound ethical analysis and policy formulation, as described in the teacher's narrative. In addition, as in the first module, the writers and the advisory board must struggle with the choice of traits addressed in the new module. For example, the new module proposes that students address "criminal genes" and "genes for mental illness." Given that the genetic contributions to such characteristics are unclear -- and may remain so even in 2015, the time of the hypothetical activities -- should we risk conveying a deterministic misconception to our students? On the other hand, we used such traits in the first module precisely to involve students in discussions of the relative contributions of genes and environment to complex traits and of the pitfalls of determinism and reductionism. This issue -- the choice of traits for the new module -- requires additional discussion.

V. POSSIBLE EXTENSION ACTIVITIES

In biology:

- students look at sequence differences among four or five different organisms and, given information on mutation rates, construct simple phylogenetic trees that propose possible phylogenetic relationships and times of evolutionary divergence

- students analyze and discuss a series of human pedigrees that illustrate a number of ethical questions having to do with the inheritance of disease-related genes

- students search for sequence similarities among various organisms and attempt to determine whether similar sequences code for proteins that have the same possible function

- students search for putative genes in a sequence database using concepts such as open reading frames

- students use simple telecommunications software to access on-line information about the cystic fibrosis gene

In literature:

- students read related literature and discuss in the light of their knowledge of the HGP (e.g., *The Scarlet Letter, The Fall of the House of Usher, Brave New World*)

In social studies:

- students discuss international collaboration in the HGP, including the problems posed by differences in patent laws from one country to the next

- students consider eugenics as an historical problem (e.g., talk about the

importance of Jesse Owens)

In computer science:

- students design and construct a database

In music and art:

- students ask questions about the nature of genes for musical and artistic ability (perhaps they complete a study of a famous musical family)

VI. RESOURCES FOR WRITERS AND PROJECT STAFF

Prototype database software

Institute of Medicine report on genetic testing

Smart Card

Office of Technology Assessment report on genetics and the courtroom

Access to OMIM on-line

Information on the MIB

Materials from the Midwest Bioethics Center, Kansas City, MO

APPENDIX D

Evaluation Instrument for the Conceptual Framework

REVIEW OF CONTENT AND PEDAGOGY
Advisory Committee Summary
Mapping and Sequencing the Human Genome: Science, Ethics, and Public Policy--Development and Distribution of an Instructional Module for the High School Biology Classroom

1. Reviewer_____

2. Date_____

3. Are the major concepts accurate?

very accurate	generally accurate	generally inaccurate	completely inaccurate

4. Are the major concepts appropriate for first-year biology courses?

very appropriate	somewhat appropriate	inappropriate

5. Do the proposed teacher background materials include the major scientific, ethical, and public policy issues raised by the Human Genome Project?

all	most	some	none

6. Are the proposed classroom activities at the appropriate level for first-year biology courses?

too difficult	about right	too easy

7. Is the proposed treatment of ethics and public policy appropriate for first-year biology courses?

very appropriate	somewhat appropriate	inappropriate

Please include specific comments on one of the copies of the summary enclosed.

APPENDIX E

Summary of the Review of the Conceptual Framework

REVIEW OF CONTENT AND PEDAGOGY
Advisory Committee Summary
The Human Genome Project:
Information Management, Access, and Regulation
Summary of Reviewers' Comments

This summary is based on the responses of 13 (out of 34) reviewers. Responses are reported as mean values.

1. Reviewer: _____

2. Representing: ASHG _____ CORN _____ NSGC _____

3. Date: _____

Please include additional specific comments on one of the copies of the enclosed summary. Thank you for your help.

Section I. Major Concepts

Indicate in the grid below whether the treatment of major concepts dealing with *biology*, *databases*, and *ethics and public policy* are *accurate*, *appropriate*, and *sufficient*.

Please use the following point values:

3 = very accurate, very appropriate, sufficient
2 = somewhat accurate, somewhat appropriate, somewhat sufficient
1 = inaccurate, inappropriate, insufficient

Major Concepts	Accurate	Appropriate	Sufficient
Biology	2.83	2.55	2.82
Databases	2.83	2.91	2.64
Ethics & Public Policy	2.92	2.82	2.73

Section II. Curricular and Classroom Considerations

Are there other issues that you think might arise with respect to curricular and

classroom considerations? (This listing is inclusive of all written responses received as of this date.)

1. Cytogenetics - Chromosomal disorders

2. What part of the introductory biology course is replaced by this module?

3. Might it be possible to utilize local library computer resources in areas where computer to student ratio is low?

4. Discussion of information related to *eugenics* is too limited. (These are likely to become public policy issues in terms of use of information.) Needs to be more specifically identified as such.

5. Also may be useful to incorporate current events.

6. This project will be very difficult in classes that do not have a sufficient number of computers for all students to use.

7. I think the hand search for similarities in the database is important, but needs to be streamlined. I suspect the students will consider looking for two, then three, then four base sequences tedious.

8. Population genetics in relationship to forensics and its use in the courtroom.

9. (I am) Worried that (there is) too much material for five days. Would like more discussion on informed consent (good way to bring in cultural sensitivity). Would like some mention of genetic health care providers, particularly genetic counselors. Seems all testing discussed in module is available without any thought to obtaining genetic counseling before testing.

10. It seems to me that there will be little time to orient students to the computer. Are you assuming that they will have some computer experience prior to starting this module? I am also concerned that some teachers will not be able to use this module because they do not have access to computers, or experience using them.

11. If the original module can be replaced with this one, the focus is too narrow. If they are to be used together, it is going to take too much time. Teachers say there is already too much to cover in biology. It may be helpful to teachers to suggest how they can use <u>parts</u> of both modules.

12. The relationship of research data to clinical applications (should be examined).

13. (Include) how access to data is decided and made possible--choices of software, etc.

14. Major concepts in "Biology" are sufficient if they follow basic genetics and the other module.

15. Use the model from the last module (in presenting) "Ethics and Public Policy" concepts.

16. Make it clear what's currently used and the idea of what's coming re: databases, access, etc.!! Give students sense of progress (rapid, changing, who shapes this, etc.).

Section III. Teacher Narrative

Please indicate in the grid below whether the proposed background materials dealing with *science and technology* and *ethics and public policy* are *accurate*, *appropriate*, and *sufficient*.

Please use the following point values:

3 = very accurate, very appropriate, sufficient
2 = somewhat accurate, somewhat appropriate, somewhat sufficient
1 = inaccurate, inappropriate, insufficient

Background Materials	Accurate	Appropriate	Sufficient
Science & Technology	2.82	2.91	2.73
Ethics & Public Policy	2.82	2.82	2.73

Section IV. Classroom Instruction

1. Is the proposed treatment of the *structure* of genome databases appropriate for first-year biology courses? (mean = 2.82)

very appropriate	somewhat appropriate	inappropriate
3 **(2.82)**	2	1

2. Is the proposed treatment of the *uses* of genome databases appropriate for first-year biology courses? (mean = 2.83)

very appropriate	somewhat appropriate	inappropriate
3 **(2.83)**	2	1

3. Will the proposed activities allow students to use the databases we provide in a manner that promotes understanding of informatics and the Human Genome Project? (mean = 2.73)

very likely	somewhat likely	unlikely
3 **(2.73)**	2	1

4. Do the models and simulations proposed in these activities provide accurate representations of issues related to storage and use of data generated by the Human Genome Project? (mean = 2.73)

very accurate	somewhat accurate	inaccurate
3 **(2.73)**	2	1

Selected Comments from Reviewers:
(As written on Summary copies)

Intro: "Prerequisite: Computer knowledge both for students and teachers. (I assume computer literacy is now mandatory in high schools.)"

p.2: "I question whether it is realistic to think students will be able to master these concepts and the information-processing technologies outlined in the grant."

p.7: #2. "Wouldn't this module be more appropriately targeted for computer science teachers who could do some team teaching with biology teachers?"

"Abbreviations are used throughout without definition."

"(Use portions of the first module as necessary, especially those portions that address the basics of ethics and ethical analysis.) This will be especially useful for teachers who use both modules--hone their "ethics" skills and confidence."

p.8: #11. "Most that I saw related to cystic fibrosis. Be sure to include other genes also."

p.10: 2A. "August '93 issue of *Journal of NIH Research* will have updated schematic (color) of all mapped genes!"

p.13: E. "Will they have time to cover this? Rather complex."

p.14: F. "What is meant by "fair" and "unfair" discrimination?"

"Ethical principles: informed consent and informed refusal."

p.22: "What are the differences between DNA fingerprinting and DNA analysis? (also on p. 26, Ex.3)"

p.26: "This class is certainly unethical! Do you really want to present this classroom model?"

p.29: (Add to Student Objectives): "How to get ideas to policymakers (Congress, etc.)."

Selected Comments from Tom Murray:

p.3: A suggestion for a metaphor (on how to present triplets, genes, sequencing, etc.).

p.23: Ex.3. "I don't think this sort of simplistic behavioral genetics should be in the unit."

p.30: "This is right on target!" (Should we risk conveying a deterministic misconception to our students?)

"Overall, this plan looks to be very creative, complete and sound. Typical BSCS work!"

Selected Comments from James Sikela:

p.5: B.10. Regarding the second statement, in the U.S., is not insurance eligibility *central* to "health and medicine"? The second statement is somewhat vague in that no examples or specifics are given.

p.7: 3. Prior instruction should cover gene structure, e.g., introns and exons, ORFs, to allow students to better distinguish features of the sequences that they will be dealing with.

General impressions:

"On the whole, I think the module is very well done and will be quite valuable to students. The few specific comments I have are as follows:

(Excepts)

1) Some additional emphasis could be placed on parallels already familiar to students. For example:

> Examples of 'screen-out' questions from job applications...Credit histories...Medical histories...Criminal records...Contemporary examples of biological "discrimination" among employers and insurers...Examples of stigma attached to entire groups because a subset of the group contracts a a disease...Historical contexts for reaction to disease...

2) There is not much discussion on the rationale for collecting genetic information. It could be made clearer that gene sequences and map information can lead not only to diagnosis and a challenge to personal and societal ethics, but to treatment as well. There could be more emphasis on the benefits of prenatal or presymptomatic diagnosis and possible therapies that may be developed as a result of the genome project.

4) The concept of human intervention in the human evolutionary process is central to the ethical issues that relate to the HGP...a fundamental assumption underlying the efforts of the HGP is that increased knowledge of human heredity will be more useful to human well-being than if such knowledge was not known."

APPENDIX F

Service Contract with Learning Systems Consultants, Inc.

July 6, 1993

Service Contract

This agreement is made this sixth day of July, 1993, between BSCS, hereafter **BSCS**, and Learning Systems Consultants, Inc., hereafter **LSCI**.

Base Agreement:

In consideration of the payment of FIVE HUNDRED DOLLARS ($500.00) payable in advance by BSCS to LSCI, and in consideration of the mutual promises, covenants and agreements hereafter set forth, and for other good and valuable consideration paid by each party to the other, IT IS AGREED:

LSCI will provide 10 hours of the following services:

Analysis of proposed software:

The Human Genome Database software (hereafter HGD) is a database-like program that will support student inquiries upon a limited size and scope "database" containing fictional genetic information. LSCI will work directly with BSCS to identify the detailed requirements of the product. The results of the Analysis will be published by LSCI in the form of a data-flow diagram and narrative specifications document.

Initial Product User-Interface Design/Prototype Development:

Based upon the results of the Analysis (above) LSCI will develop prototypical software that will allow BSCS to "test-drive" proposed user-interfaces for HGD. It is understood that the prototype will not be fully functional, rather it will simply allow BSCS to demonstrate the concept and user-interface of the proposed HGD. It will function only to provide a simulation of what functionality the completed production version might in part perform. The results of the Initial Product User-Interface Design/Prototype Development will be published by LSCI in the form of "screen shots" and transfer of ownership of the prototype to BSCS. The HGD prototype code developed by LSCI will become the sole property of BSCS at completion of this contract.

Product Development Bid:

Based upon the results of the Analysis and Initial Product User-Interface Design/Prototype Development (above), and the feedback/recommended modifications to the analysis and design specifications provided and documented by BSCS, LSCI will produce a bid for completion of the finished production version of the HGD product, based upon time and resources required. LSCI realizes that BSCS is required, under the terms of the controlling grant from the department of energy, to secure at least three bids for contract work and to make decisions in accordance with federal grant requirements.

Provision for Expanded Scope:

LSCI will make every effort to complete the project within the provisions of the Base Agreement. However, upon completion of the Analysis phase (above) and review of potential screen designs for the prototype, if the size and scope of the project and/or the hours required to refine requirements have substantially increased beyond the original ten (10) hours provided for by the Base Agreement, LSCI may request additional compensation of up to FIVE HUNDRED DOLLARS ($500.00), for up to an additional ten (10) hours of like support. This provision is only in effect provided BSCS approves the additional hours in order to produce a high quality prototype that accurately reflects the product direction and intent.

Signed this ____ day of _____, 1993

In the presence of:

_____ _____

Witness Joe McInerney
Director-BSCS

_____ _____

Witness Jeff Thomas
President-Learning Systems Consultants, Inc.

APPENDIX G

Description of the Prototype Databases

Description of the Prototype Databases

I. Summary of the Basic Design Features

The prototype databases were designed for use on any Macintosh. The project staff anticipate that the full working databases also will be developed for the Macintosh and that they will be modified to run on MS-DOS and Apple machines only after the first field test and after completion of subsequent revisions of the activities and the database.

The BSCS and LSCI staff have designed the databases to meet several key criteria. These include 1) the requirement that the databases be easy to use, 2) the requirement that the program file be small (to reduce downloading time and cost), and 3) the need for the databases to allow specific functionalities. For example, the prototype was designed with the expectation that the final product will be fully self-contained, and require no additional software (e.g., word processors or database programs) to run. Likewise, the design parameters call for the use of self-extracting compressed files, which should reduce downloading time significantly. Finally, BSCS and LSCI staff have made both design and programming decisions that have avoided including features or using language that might make the database more robust but would not add functionality important to the curriculum.

II. Summary of the Conceptual Organization

The prototype software for the project consists of two separate databases: the "National Genome Database" (NGD) and the "Local Genome Database" (LGD).

The NGD is organized as a research database, consisting of a set of GenBank-like entries. These entries include, for each of 10 human genes, such information as a 25-100 base consensus sequence, a summary of function, and other appropriate information (e.g., a list of references, and the amino acid sequence). The NGD also includes data on selected genes from other organisms. Data in the NGD are searchable by gene name, organism, and sequence, and allow users to answer such questions as "What is known about this gene?", "What genes for this organism are known?", and "In what gene or genes is this sequence found?" A screen print from the prototype showing the record for the human alkaline phosphatase gene is shown in Figure One.

The LGD is organized as a registry containing personal genomic data on a fictitious population. Each record in the LGD corresponds to one person and includes a "name/number" field for identification, fields for the base sequences for each of ten genes, and a field to indicate sex. The LGD is searchable by sample number, gene name, sequence, and sex and allows users to answer such questions as "What is the genetic profile of this individual?" and "Who are the individuals in this database who carry this sequence?" A screen print showing a sample record from the prototype LGD is shown in Figure Two.

Figure 1. Sample screen print from the NGD.

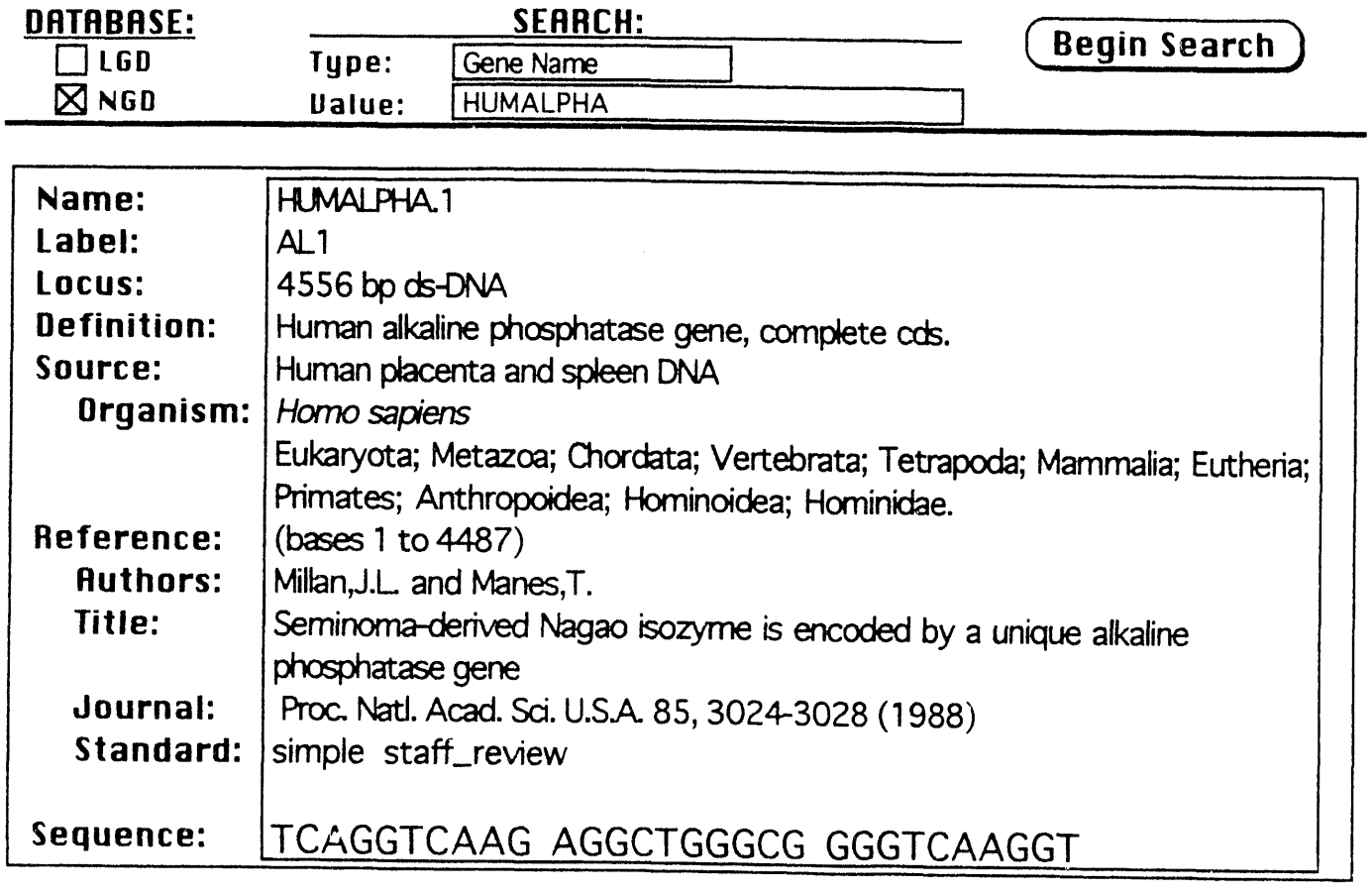
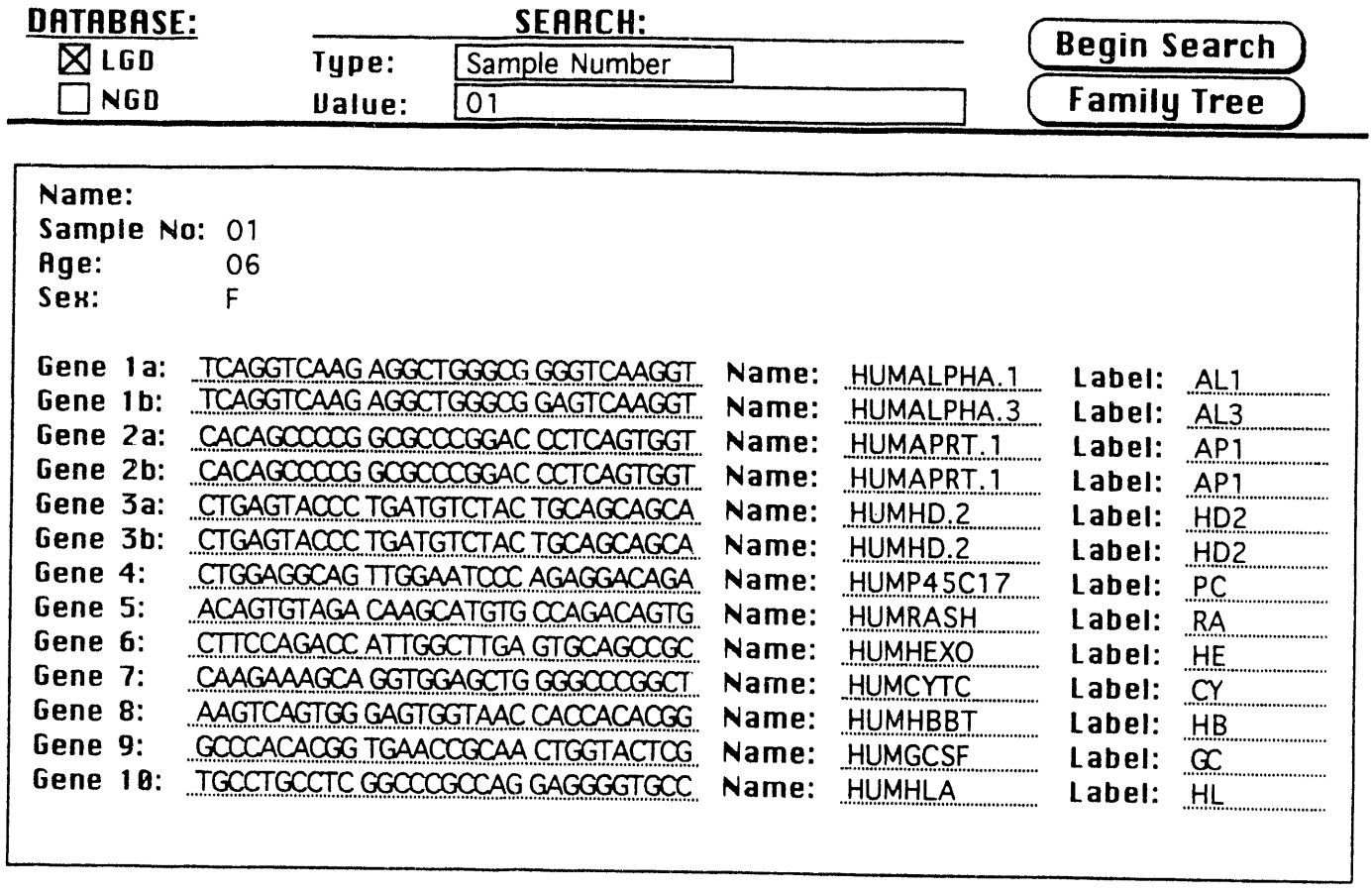


Figure 2. Sample screen print from the LGD.



APPENDIX H

Agenda for the Writers' Conference

AGENDA
of the
WRITING CONFERENCE
for
The Human Genome Project: Information Management, Access, and Regulation
19 - 30 July 1993

MONDAY, 19 JULY - Worner Center, Room 216 (Colorado College)

7:30 am	Continental breakfast (set up outside Room 216)
8:00 - 8:15 am	Welcome and introductions - logistical issues
8:15 - 8:45 am	Overview of the project - goals and objectives - intended audience - time line for materials development
8:45 - 11:00 am	Discussion of the advisory committee's framework and summary of the feedback from the external reviews - major concepts - curricular and classroom considerations - teacher's narrative - instructional activities
[9:45 - 10:00 am Break - outside Room 216]	
11:00 am - 12:00 pm	Demonstration and discussion of the sample database
12:00 - 1:00 pm	Lunch (set up outside Room 216)
1:00 - 2:00 pm	Discussion of revisions to the framework
2:00 - 3:00 pm	Discussion of writing tasks - teacher's narrative - instructional activities - assessment tools

3:00 - 3:45 pm Writing-team assignments; procedures for manuscript preparation; review of schedule for the conference

3:45 - 5:00 pm Tour of BSCS, writing facilities

TUESDAY, 20 JULY - BSCS Headquarters

8:00 am - 5:00 pm Writing

6:00 pm - 7:00 pm BSCS Sky Sox Game tailgate party

7:05 pm Sky Sox Game

WEDNESDAY, 21 JULY - BSCS Headquarters

8:00 am - 5:00 pm Writing

THURSDAY, 22 JULY - BSCS Headquarters

8:00 - 10:00 am Writing

10:00 am - 12:00 pm Discussion of group progress (BSCS conference room, 3rd floor)

- teacher's narrative
- instructional activities
- database

1:30 - 5:00 pm Writing

FRIDAY, 23 JULY - BSCS Headquarters

8:00 am - 5:00 pm Writing

6:30 pm Dinner at Zeb's

SATURDAY, 24 JULY AND SUNDAY, 25 JULY

Rest and relaxation

MONDAY, 26 JULY

8:00 am - 5:00 pm Writing

TUESDAY, 27 JULY

8:00 am - 5:00 pm Writing

WEDNESDAY, 28 JULY

8:00 - 10:00 am Discussion of group progress

- teacher's narrative
- instructional activities
- database

10:00 am - 5:00 pm Writing

THURSDAY, 29 JULY

8:00 am - 5:00 pm Writing

FRIDAY, 30 JULY

8:00 am - 12:00 pm Writing
12:00 - 1:30 pm Lunch (set up outside Room 216, Worner Center, Colorado College)
1:30 - 3:00 pm Summary of group progress
3:00 - 3:15 pm Break (set up outside Room 216)
3:15 - 4:30 pm General discussion

- project evaluation (including field-test questions about the database)
- additional writing assignments
- teacher orientation
- miscellany

4:30 pm Adjourn

APPENDIX I

List of Participants for the Writers' Conference

The Human Genome Project: Information Management, Access, and Regulation

Participant List

Mary Ann Cutter, PhD
University of Colorado at Colorado Springs
Department of Philosophy
1420 Austin Bluffs Parkway
P.O. Box 7150
Colorado Springs, CO 80933-7150
school (719) 593-3340

Edward Drexler
Pius XI High School
135 N. 76th St.
Milwaukee, WI 53213
home (414) 258-9190

Kay Gottesman
Manager, User Services
Human Genome Data Base
2024 E. Monument St.
Baltimore, MD 21205
office (410) 955-0289
FAX (410) 614-0434
E-MAIL: KSG@LIBRARY.WELCH.JHU.EDU

Laurence B. McCullough, PhD
Center for Ethics, Medicine, and Public Issues
Baylor College of Medicine
1 Baylor Plaza
Houston, TX 77030
(713) 798-6290

Richard J. Mural, PhD
Senior Staff Member/Biology Division
Oak Ridge National Laboratory
P.O. Box 2009
Oak Ridge, TN 37831
Lab phone: (615) 576-2938
Div. office phone: (615) 574-0212
FAX: (615) 574-1274
EMAIL: M9L@stc10.ctd.ornl.gov

Jeffrey C. Murray, MD
Division of Medical Genetics
Department of Pediatrics
University of Iowa Hospitals & Clinics
Iowa City, IA 52242
office (319) 356-3508
FAX (319) 356-3347
EMAIL: jmurray@umaxc.weeg.uiowa.edu

John Zola
125 S. 34th
Boulder, CO 80303
home (303) 494-7313

Database Consultant

Jeff Thomas
Learning Systems Consultants, Inc.
824 Chelsea Ct.
Colorado Springs, CO 80918
(719) 599-7104
FAX - same number

BSCS Staff

Joe McInerney, principal investigator
EMAIL: JMCINERNEY@CC.COLORADO.EDU

Lynda Micikas, project director

Phil Goulding, technical specialist
EMAIL: PGOULDING@CC.COLORADO.EDU

Randy Backe, evaluator

Pamela Van Scotter, editor

Dee Miller, secretary
BSCS Phone: (719) 578-1136
FAX: (719) 578-9126

APPENDIX J

Preliminary Outline of the Experimental Materials

Outline of the Experimental Materials

Foreword

Section I: What Is the Human Genome Project?

Section II: The Science and Informatics of the Human Genome Project

Section III: Ethical and Public-Policy Dimensions of Research Databases and Registries

Glossary

References

Classroom Management

Annotated Student Activities

- Activity 1
- Activity 2
- Activity 3
- Activity 4
- Activity 5

Student Activity Copymasters

Detailed Outline of the Teacher Narrative

Section I: What is the Human Genome Project?

- A. The History, Organization, and Funding of the Human Genome Project
- B. Who Works on the Human Genome Project?
- C. What Will We Learn From the Human Genome Project?
  1. Insights into basic biology
  2. Diagnosis, prediction, and treatment of disorders
  3. Methods for storing and analyzing data
  4. Development of new technologies
  5. Effects on the practice of science

Section II: The Science and Informatics of the Human Genome Project

- A. The Context of the Human Genome Project
- B. The Objectives of the Human Genome Project
- C. Information About Where Genes Are Located on Chromosomes
  1. Genetic linkage maps
  2. Techniques used to generate linkage maps
  3. Physical maps
  4. Techniques used to generate physical maps
- D. Information About the Nucleotide Sequences of Genes
- E. The Role of Electronic Databases in the Human Genome Project
  1. Electronic databases
  2. The storage of genomic information in electronic databases

3. Research databases

4. Map data

5. Sequence data

6. Clinical data

7. Other research databases

8. Registry databases

F. The Use of Research Databases in Biology and Medicine

1. Locating genes on chromosomes

2. Recognizing genes by their internal organization

3. Predicting the functions of gene products

4. Determining evolutionary relationships

5. Understanding the clinical implications of genetic conditions

G. The Use of Registry Databases in Biology and Medicine

H. Scientific, Technological, and Ethical Issues Associated with Genomic Databases

1. Problems of database design

2. Problems of data accuracy

3. Problems of data analysis

4. Problems of privacy

5. Problems of misuse

I. The Future of Genomic Databases

1. Importance of genomic databases to the HGP

2. Standardization across databases

3. Creation of an integrated system

4. Issues of storage and speed

5. Security

6. Careers in informatics

J. Concerns About Genomic Databases

Section III: Ethical and Public-Policy Dimensions of Research Databases and Registries

A. Ethical and Public-Policy Questions Concerning Access To and Use of Information in Research Databases and Registries

1. Issues concerning research databases

- the cost of research databases
- access to research databases by scientists and others from countries that did not contribute to the cost of producing the databases
- use of research database information to define "normality"
- use of a generalized genome in a multiracial society and world
- self-counseling

2. Issues concerning registries

- informed consent
- control of access to personal genomic data
- restrictions on the use of personal genomic data

B. Ethics

1. The features of ethics

2. How to talk about interests

3. The role of argument in ethical inquiry

4. Recurrent ethical concepts related to research databases and registries

- privacy
- confidentiality
- autonomy
- paternalism
- discrimination

C. Public Policy

1. *De facto* public policy
2. *De jure* public policy
3. Conditions for *de jure* public policy

D. Teaching Ethics and Public Policy in the Classroom

Summary of Proposed Activities

Activity 1

Focus

This activity introduces students to the model databases and illustrates 1) the differences between research databases and registries and 2) the usefulness of electronic storage and search techniques in dealing with genomic data.

Summary

In an opening activity designed to engage the students' attention and interest, the instructor distributes duplicated sheets containing hundreds of bases of sequence data, and asks students to perform a series of simple searches of the data. For example, she may ask students to find and to mark all the "Cs" that they can in a period of one minute, to find all the "ACGs" that they can in a second period of one minute, and, finally, to find and mark all the "GCCTACGTAACGGTAAGs" that they can within the third period of one minute. This "engage" activity ends with the teacher observing that the model sequence only consists of several hundred bases (in contrast to the full human sequence that consists of some 3 billion bases), and that information about base sequence is only one of the kinds of genomic information that scientists must collect and analyze as part of the HGP.

The instructor explains to the students that in the next few days they will explore some of the advantages as well as some of the potential problems associated with the electronic storage and manipulation of genomic data. She divides the class into small groups for work on the computer, and gives each student a preassigned identity matching that of someone whose name and genomic profile are stored in a registry database that has been created for a fictitious town. She asks the students to use the model research (NGD) and registry (LGD) databases to answer the following questions about "themselves":

- What is my personal profile (e.g., how old am I, what is the state of my general health, who are my parents and my siblings)?
- What is my genetic profile (e.g., what are my base sequences for each of the genes for which data are stored in the LGD)?
- What are the likely genetic and/or medical implications of these sequences?

To answer these questions, students will have to search the LGD for their personal records, and then search the NGD (the research database) for information about one or more of their sequences (this will depend on the time available).

The homework assignment for students will be to: 1) draw a small pedigree showing the correct relationships among members of their immediate family; 2) identify themselves on the pedigree and indicate their genotype using proper genetic symbols (found in the NGD); and 3) write a short paragraph describing the genetic and/or medical significance of one of the alleles that they carry.

Student Objectives

As the students complete this activity, they should:

- understand that the volume and complexity of genomic data are such that electronic databases are required to store and search them effectively;
- search a registry database for personal genomic data;
- search a research database for information that explains and expands upon the personal data;
- construct a pedigree on the basis of family relationships reported in the registry;
- use proper genetic notation to indicate the genotypes of individuals in the pedigree; and
- use information gained from the research database to explain the genetic and/or medical implications of an DNA sequence of unknown origin.

Science Process Skills

- Observing
- Comparing
- Gathering data
- Analyzing data
- Synthesizing information and knowledge
- Communicating (orally and in writing)

Activity 2

Focus

Activity 2 allows the students to see how scientists use both research databases and registries to test hypotheses about genetic data.

Summary

At the start of the activity, students work in their original groups to combine their individual family pedigrees into a consolidated pedigree, in effect, generating a hypothesis about the structure of their "extended" family. When students access the LGD to verify their hypothesis, they discover an "outlier," i.e., one person in their family whose genetic profile is not consistent with his or her position in the pedigree.

The instructor asks the students to develop a set of hypotheses to explain the apparent genetic discrepancy (e.g., the person was adopted, one of the person's parents is not as reported, sample errors, data-entry errors, malicious mischief), and then to use the LGD and the NGD to narrow the list of possible explanations as much as possible. In one case, students discover that the reported sequence is actually not a human sequence, but a canine sequence; they conclude that there likely was a sampling error. In another case, students determine that the individual's genotype is inconsistent only with his mother's genotype, and conclude that this may have been a case of *in vitro* fertilization using a donor egg. In a third case, students discover that the only inconsistent sequence is not contained in any other place in the databases (either the LGD or the NGD), and they discuss what further experimental steps they would take to determine whether this was a data entry error or whether they had discovered a new allele for the gene in question.

The day's work ends with the teacher asking students to identify some of the possible positive effects of our increased knowledge of the human genome (e.g., a better understanding of our individual genetic profiles can help us make life-style decisions that can improve health), as well as some of the possible negative effects (e.g., inaccurate or out-of-date information can cause us undue concern and anxiety).

Student Objectives

As the students complete this activity, they should:

- combine individual pedigrees into a consolidated pedigree that correctly describes a set of reported family relationships;
- identify cases in which genetic data are not consistent with reported relationships;
- develop a set of hypotheses to explain discrepancies between genetic data and reported family data;

- search a registry database for information that will distinguish among such hypotheses;
- search a research database for information that will distinguish among such hypotheses; and
- observe and analyze a concrete example of the ways in which genomic information stored in electronic databases can be used to answer questions of general and personal interest.

Science Process Skills

- Gathering data
- Analyzing data
- Generating hypotheses
- Predicting
- Inferring
- Communicating (orally and in writing)

Activity 3

Focus

This activity offers students the opportunity to use the NGD to explore a newly discovered genetic phenomenon (genetic anticipation), while also raising questions about the privacy of genetic information, and about possible just and unjust discrimination that can result from unregulated to access and/or uninformed use of genetic information.

Summary

As the activity opens, the teacher announces that a group of scientists working in a national genome center have added some new information to the NGD concerning a previously poorly understood sequence from the X-chromosome. Unfortunately, however, an article in the local newspaper about the discovery has so upset some of the parents in the fictitious town that they did some investigating in the LGD and now have told their teenage sons that they may not date or have anything to do with a particular girl. The parents have refused to explain their decision to the boys, or to the girl's parents.

The teacher asks the students to use the LGD and the NGD to determine why the parents might have made such a decision. To answer the question, students must access the girl's personal file in the LGD, and research the genetic and medical implications of each sequence listed. They discover that her genetic profile is relatively unremarkable, except that the new data show that her repeat length -- 54 -- in the newly identified fragile X sequence places her either at a high normal or a low premutation position on the spectrum of possible genotypes. (The trinucleotide repeats that occur in the fragile X gene are polymorphic in the normal population with ranges found from 6 repeats to 54; in fragile X premutations, the smallest number of repeats reported is 52[1].)

As the period closes, the teacher helps the students summarize their findings (e.g., the LGD shows no history of mental retardation in the girl's family, but her repeat length is high) and asks them whether they think the evidence is sufficient to conclude that the girl is at risk for bearing a mentally retarded child, and what the girl might do to get more insight into her genetic situation. If students begin to raise questions about the fairness of the parents' decision or about whether anyone should have had access to these data anyway, the instructor allows them to talk, but does not attempt (in the little time remaining) to direct a full ethical analysis with them. Instead, she leaves them with the following open question "Now that you've spent three days searching freely through these databases, do you see any reasons why allowing completely

[1]Nelson, David L. Fragile X syndrome: Review and current status. *Growth: Genetics & Hormones*. Vol 9, No. 2, June 1993.

unrestricted access to all genetic data might not be in the public's best interests?" The teacher explains that Activities 4 and 5 will give students the opportunity to analyze an actual case having to do with the privacy of genetic information stored in a registry and that at the end of Activity 5, the students will be asked to reconsider the details of Activity 3 in light of what they have learned.

The homework assignment in preparation for Activity 4 is to read the following article "An Ethical Quandary: The French Glaucoma Case" (*Science*, 19 April 1991).

Student Objectives

As the students complete this activity, they should:

- search a registry database for information about the genetic profile of a hypothetical individual;
- search a research database for information that would help explain that genetic profile;
- use the information in the research database to explain the phenomenon of genetic anticipation as it relates to a specific hypothetical situation;
- evaluate the significance of the genetic data with respect to the reproductive prospects for a hypothetical individual; and
- ask questions about the privacy of genetic information and about the possible consequences of misinterpretation and/or misuse of such data.

Science Process Skills

- Gathering data
- Analyzing data
- Evaluating data

Ethics Process Skills

- Recognizing ethical issues
- Gathering information
- Organizing information
- Synthesizing information and knowledge

AN ETHICAL QUANDARY:
THE FRENCH GLAUCOMA CASE

DIRECTIONS

Read the following actual case study about hereditary juvenile glaucoma in France. This case study appeared in 1991 in *Science*[1], one of the most prestigious international scientific journals. When you have read the case study, answer the question that appears in the last paragraph. Feel free to underline and highlight what you think are important aspects of this case.

> Paris--A team of researchers sifting through 5 centuries of French village records for patterns of mental illness has instead turned up an astonishing pattern of blindness caused by hereditary juvenile glaucoma--a pattern that goes all the way back to a single couple living in a village in Brittany in the 15th century. The researchers have since traced no fewer than 30,000 living Frenchmen and Frenchwomen who are descended from that couple, and they have found that more than half of all reported French cases of juvenile glaucoma have occurred in people in that direct lineage.
>
> The researchers, from the Institut National d'Etudes Demographique (INED), were elated--treated early with drugs or surgery, this form of glaucoma can be arrested; blindness occurs only in untreated sufferers. So INED's data could be invaluable in pinpointing families at risk and ensuring that they get early treatment. But then came a revelation: French privacy law, designed to protect at almost any cost the privacy of the French citizenry, would prevent any such use of the information.
>
> "I know the names of the people, often young ones, who risk becoming blind tomorrow, but I cannot alert them," says André Chaventré, director of INED's Department of Anthropology and Genetic Demography, who led the team that traced the genealogy of the disease. And Chaventré isn't the only one who's incensed. Claude Evin, minister of Social Affairs and Solidarity, recently announced the results of the INED study at a medical ethics conference and has since done his best to get the privacy rules changed.
>
> The identification of potential bearers of the putative glaucoma gene is the fortuitous result of a study Chaventré started 3 years ago with psychiatrist Edouard Zarifian of the Caen University Hospital. They were trying to trace the genetic pattern of manic depression and soon realized that there was a strong, but

[1]Dorozynski, A. (1991). Privacy rules blindside French glaucoma effort. *Science* 252:369-370.

so far unexplained, statistical link between this disease and a common variety of congenital juvenile glaucoma known as open-angle glaucoma. The disease is insidious: the patient, often a child, does not become conscious of the disease until vision is affected, but by that time a large proportion of optic fibers are irreversibly damaged.

Chaventré came across a 1979 medical thesis reporting a high incidence of juvenile glaucoma in the Nord-Pas-de-Calais region, near the English Channel, and quickly recognized it had a familial pattern. Chaventré contacted ophthalmologists in Lille and Paris and established a protocol to trace the genealogy of manic depression, glaucoma, and diabetes, which is known to be associated with glaucoma. The study was extended to relatives of glaucoma patients, who were given an ophthalmologic examination, glaucoma tests, and, whenever possible, psychiatric evaluation.

INED researchers assembled bits and pieces of a genealogic tree, using town and village records, often kept in several copies by the French administration. Posted on a wall, the tree was several tens of meters long. Computer analysis unequivocally pointed to a single couple, who died in 1495 in a small hamlet near the village of Wierre-Effroy in the department of Pas-de-Calais, as the original source of the disease. (An 11th-century chapel in Wierre-Effroy, dedicated to Sainte Godeleine, contains a cistern filled with water that was believed to cure blindness: even today, pilgrims gather there every year in July to pray for the healing of the blind. "This," says Chaventré, "is not a coincidence.")

From this 15th-century couple, the gene spread rapidly throughout the region and the country. "This can go very fast," says Chaventré. "We have found records of affected parents who had as many as 18 children." The data are now coded and stored on a computer in the INED building in Paris. And if the Commission Nationale d'Informatique et des Libertés (CNIL) gets its way, that's where they will stay.

In 1988 Chaventré consulted CNIL, which was created in 1978 to protect individuals from potential abuses of computerized data, about a plan to inform physicians of the names of at-risk individuals living in their area. Physicians would then be able to keep a close watch on specific patients and, when necessary, recommend an examination in ophthalmology departments of designated hospitals. The CNIL cut the ground out from under the plan, however, by ruling that it would be fine for INED to tell physicians to keep an eye out for juvenile glaucoma among their patients, but it couldn't mention the names of any individuals. INED, it said, can alert physicians only to the symptoms and hereditary nature of the disease.

Chaventré objects that alerting physicians without telling them which patients are at risk would be ineffective, and that a national screening campaign would overwhelm specialized centers. "Giving physicians the names of individuals

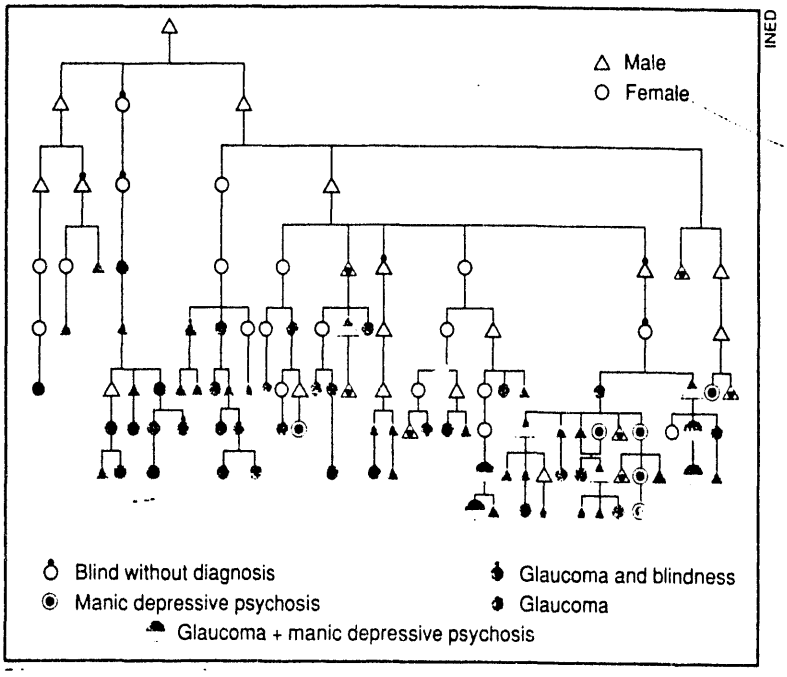
registered in their neighborhood, who are on the INED list would be far more efficient," he says. But Vulliet Tavernier, an official at CNIL, counters that distributing a list of individuals obtained by a genealogic study would constitute an authoritarian public health measure that would infringe on individual liberty and privacy. CNIL is concerned that circulating the names of potential carriers of genes predisposing to diseases might lead to discrimination in hiring or insurance.

CNIL bases its legal case on a 1978 law that states that individuals about whom information is collected must know how the information will be used. The law specifically notes that "even in the domain of medical research, such information can, in certain cases, cause prejudice to a patient because it informs him he is affected by a severe disease." Although a proposal was floated in 1989 to change this legislation to permit some types of data to be released to protect public health, it was rejected because "they did not provide for a satisfactory equilibrium between the interests of public health, the respect of fundamental liberties, and the rights of men, notably the right to respect privacy," CNIL president Jacques Fauvet wrote at the time.

Meanwhile, Evin, whose jurisdiction includes health, has forced a public debate on the INED study. During a congress on ethics organized by the Conseil National de l'Ordre des Médecins, the French National Medical Association, last month, he said, "The use of informatics can be felt as a threat....But techniques of genealogical studies in France allow the identification of thousands of persons at risk for certain diseases that can perfectly well be prevented."

Should CNIL allow the release of th of families and individuals at risk of hereditary juvenile glaucoma?

Glaucoma genealogy. Fragment of family tree of one 15th-century couple shows transmission of juvenile glaucoma and manic depression.



Activity 4

Focus

In this activity, students will use the skills of ethical reasoning to analyze the ethical issues raised by an actual case that occurred in France in 1991. This exercise sets the stage for Activity 5, in which students consider the public-policy implications of the same case.

Summary

The teacher introduces the activity by reminding the students that the case in question is an actual situation that has faced French society, scientists, and policy makers, and that the steps that the investigators took in this case are similar to many of those undertaken by students during the first few days of this unit (e.g., constructing a pedigree by identifying family relationships and determining genetic variability and analyzing its inheritance patterns).

The teacher then leads the class through an analysis of the ethical dimensions of the French glaucoma case. She begins by asking the class how many students answered "yes" to the homework question ("Should CNIL allow the release of the names of families and individuals at risk of hereditary juvenile glaucoma?") and how many said "no." Either in small groups (option A) or as the whole class (option B), students identify and clarify arguments to support both the "yes" and "no" positions and the teacher lists these reasons on the chalkboard. Once the lists are complete, the teacher asks the students to identify what they consider to be the most compelling reasons on each side of the issue, and invites individual students to explain their choices. The teacher then asks whether the two sets of reasons are equally compelling in response to the case study, and if so, whether this is a satisfactory outcome to their discussions.

The pedagogical goal is to help students see that sometimes competing ethical reasons emerge and that it is not always possible to decide once and for all between them. This outcome, when it occurs, makes ethical decision making difficult, but people who disagree should respect those with whom they disagree and not think them to be persons of bad will.

The teacher concludes the lesson by asking the class whether they can think of any other situations in which access to personal genetic information in a registry can create ethical dilemmas. As homework, she asks the class to consider the work they have done during the week and to come to class on the next day prepared to defend the existing French policy on release of personal medical data or to suggest an alternative policy.

Student Objectives

As the students complete this activity, they should:

- use the skills of gathering information, evaluating information, and making and analyzing arguments as tools for ethical inquiry;

- evaluate ethical issues related to registry databases;
- take and explain a position on whether respect for autonomy or the pursuit of valued consequences is the more important ethical consideration regarding the case study; and
- appreciate that other students will have well-argued, sometimes quite divergent, alternative views with respect to these issues.

Science Process Skills

- Gathering data
- Describing
- Inferring
- Synthesizing information and knowledge
- Communicating (orally and in writing)

Ethics Process Skills

- Gathering information
- Evaluating information
- Making arguments
- Analyzing arguments
- Recognizing and accepting disagreement

Activity 5

Focus

In this activity, students will consider the public-policy dimensions of the case introduced in Activity 4. Class discussion should build on the accomplishments of students during the earlier discussion of the ethical issues involved in this case and should reflect the scientific and technical knowledge that students have acquired about research databases and registries during Activities 1-3. A goal for this final activity is for students to understand that public policy, when done carefully, is a powerful form of "preventive ethics," because effective policy anticipates and addresses likely as well as unlikely ethical concerns.

Summary

The teacher will conduct the activity in two stages. First, she will divide the class into small groups of 3-5 students and ask students to compare their answers to the homework question (i.e., whether or not they agree with the French policy on release of personal medical data). After 5-10 minutes, the teacher will ask the whole class to identify the strongest arguments for maintaining the current French policy and the strongest arguments for modifying it.

In the second stage, the teacher will distribute the following handout to each student and allow the class 5-10 minutes to complete it. She then will direct a whole-class discussion focused on the following questions:

- What were their choices among the four options listed in the handout for the best policy for dealing with the French case?
- If their recommendations were implemented, what would happen in the French case?
- How might each policy meet or fail to meet the criteria of urgency and effectiveness?
- If they were an elected legislator, which of these options would they select as the one most likely to serve as public policy for dealing with requests for the release of data from genomic registries?
- If their recommendations were implemented, what would happen in each of the scenarios described on the worksheet?

In the final homework assignment associated with the module, the teacher asks students to write a brief paper that responds to the question: "What benefits and drawbacks do you see in the future as the Human Genome Project continues and as the amount of data in research databases and registries increases?"

Student Objectives

As students complete this activity, they should

- appreciate why public policy about genetic registries should favor individual autonomy (because this presumption furthers the interest in self-determination);
- appreciate why public policy about genetic registries should favor individual and societal health (because this presumption furthers the pursuit of highly valued consequences for individuals and society);
- appreciate that public policy that favors autonomy can be overridden by arguments that show that urgency and effectiveness of protecting individual and societal health can be satisfied only by overriding autonomy (in this case study, means is satisfied already and this is not an issue in class discussions); and
- explain that society can respond in public policy to ethical dilemmas and issues about genetic registries in one of four ways:
    1. by enacting laws that favor autonomy as the overriding consideration;
    2. by enacting laws that override a presumption of autonomy in favor of individual and societal health;
    3. by enacting laws that establish a process for consensus building about trade-offs on a case-by-case basis; or
    4. by not enacting laws -- a strategy that allows for ongoing ethical inquiry and public debate between policies that favor autonomy and those that favor individual and societal health.

Public Policy Process Skills

- Analyzing issues
- Evaluating issues
- Communicating (orally and in writing)

Handout #00
Determining Public Policy

Part I

There are several options for creating effective public policy to address situations such as the French case of hereditary juvenile glaucoma. There are two basic conditions for moving from an ethically justified position, such as we discussed yesterday, to public policy.

> Effectiveness means that the policy option will achieve its goal without great public opposition.
>
> Urgency means that there is immediate risk of serious, far-reaching, and irreversible harm if legislation is not enacted or law is not changed.

Below, you will find four possible general public-policy options. Please rank these options from 1-4 with 1=most preferred for this case and 4=least preferred for this case. Think about your reasons and the likely strengths and weaknesses of each option. Be prepared to discuss your ranking of the four options in terms of effectiveness and urgency.

A. There should be laws that establish a process for deciding about the release of registry data on a case-by-case basis.

B. There should be laws that guarantee an individual's absolute right to privacy. This means that no data could be released from this registry without actual informed consent of the individual.

C. There should be laws that greatly limit an individual's right to privacy in matters when release of the data would be in the best interest of that individual or of the community as a whole.

D. There should be no laws about rights to privacy or release of registry data at this time.

Part II

What would happen in each of the following mini-scenarios if your first choice above became law in the United States? Write brief ideas in the space provided.

A. The Arapaho Women's Health Clinic has a large genetic registry with data on its patients. Included in this registry are genetic data about susceptibility to hypertension during pregnancy, as well as other personal sexual information. Hypertension during pregnancy causes serious health risks for the woman and can involve life-threatening stroke. ACNE Pharmaceuticals is marketing a new drug that is particularly effective and safe in preventing hypertension during pregnancy. ACNE Pharmaceuticals has requested access to the registry to target

its marketing efforts.

B. Females who are carriers of fragile X often have learning disabilities. Amalgamated School District #2 wants access to registry data to identify and track these students.

C. Coach Leibniz, a former NBA hopeful who runs a summer basketball camp for teens, wants access to registry data to identify those middle school students who have a genetic predisposition for being tall. He believes strongly that if he works with these teens early, they might have a better chance of succeeding in their basketball careers.

APPENDIX K

Description of the Field-test Version of the Databases

Description of the Field-Test Version of the Databases

I. Summary of the Basic Design Features

BSCS and LSCI staff will create the field-test version of the databases based upon the prototype database program developed under the BSCS/LSCI Service Contract dated 10 July 1993, a list of general requirements developed by the BSCS staff, and a list of changes that resulted from the writers' conference.

Like the prototype, the field-test version will be developed for use on any Macintosh, will be designed for easy use, and will be fully self-contained (i.e., will not require any additional software to run). BSCS and LSCI will develop MS-DOS and Apple versions of the program only after the Macintosh version has been field-tested, revised, and approved in final form by the BSCS staff.

II. Summary of the Conceptual Organization

The field-test version of the software for the module will consist of the same two databases developed for the prototype: the "National Genome Database" (NGD) and the "Local Genome Database" (LGD).

The NGD will be organized as a research database containing general information and mapping, sequence, and clinical data on a variety of human genes relevant to the student activities. For example, Figure Three shows some of the information that might be included in the NGD record on the hypertension predisposition gene. The NGD also will contain information on a selected group of genes in other organisms as required by an extension exercise in molecular evolution (to be developed by BSCS staff). The NGD will be searchable by genetic trait, organism, genotype, sequence, and chromosome number. In all cases, the search results will display one record if the search argument is unique, and a list of records if the search argument is not unique. In the second case, the user will be able to select from the list to see one unique record.

The LGD will contain personal genomic data on more than 50 fictitious individuals comprising three different extended families. Figure Four shows a possible design for records in the LGD. The LGD also will contain full pedigrees for each of the extended families. Users will be able to search the LGD by sample number, sex, sequence, genotype, and name. As in the case of the NGD, a search will retrieve one record if the argument is unique, and a list of records if it is not.

Once a set of LGD or NGD records has been made active by retrieving them through a search, users also will be able to ask the computer to compare the base sequences attached to these records. Users will invoke this function by clicking on the "COMPARE" button that will become visible after a search has been completed. The COMPARE function asks the user to identify the standard against which a selected sequence (or set of sequences) will be compared. Once the user has identified the standard and the sequence(s) to be compared, the computer will ask the user to select the view required. Possible views will include a 1:1 comparison via a

partial dot-plot (i.e., a 30-cell by 30-cell box that indicates each match along the diagonal from left to right) or a 1:1 or 1:many comparison via a stacked display (i.e., a base-by-base comparison of sequences that are lined up one below the other). The COMPARE function will allow users easily to identify similarities and differences among any set of 30-base sequences, whether retrieved from the NGD, retrieved from the LGD, or entered manually.

Figure 3. Sample information for the NGD record on the hypertension predisposition gene.

GENERAL INFORMATION: HYPERTENSION

Hypertension predisposition gene: A gene was first identified in 1992 as one of many contributors to the risk for having early development of high blood pressure or hypertension. The gene interacts with many other genes and environmental factors such as high-salt diets, obesity, and smoking, and so is not the only factor involved in high blood pressure. One common genetic variant is known (3a); this variant makes up 90% of all alleles and has no associated risk. A second variant (3b) makes up 10% of alleles and contributes to a risk for early high blood pressure. Individuals with one or two copies of the 3b alleles should be particularly careful about diet, exercise, and smoking habits to minimize the risks of the associated allele. High blood pressure can begin in the teenage years and lead to heart attacks and strokes.

GENOTYPES: HYPERTENSION

Hypertension allele 3a: CAC TGG GTT CCT TGC TAT CGA TGG GTC AGC
Hypertension allele 3b: CAC TGG GTT CCT TTC TAT CGA TGG GTC AGC

The G to T difference in position 14 changes a threonine to a lysine in the hypertension allele3b protein.

MAP DATA: HYPERTENSION

Although it is likely that many genes and environmental factors contribute to blood pressure, one gene that plays a part in the development of high blood pressure has been mapped to chromosome 1 in humans. This gene is called the angiotensinogen gene, and variations in its DNA sequence are associated with the development of high blood pressure in pregnant women. The gene has been mapped to the long arm of chromosome 1, at the position called 1q42.

Figure 4. Sample information for the LGD record on Joy Major.

45.

Name: Joy Major

Sex: Female

Age: 15

Current status: Joy is a sophomore at Lincoln High School. She was so impressive as a freshman debater that she was promoted to the varsity debate team and is already the number two debater.

Parents names: Laura and David Major

Siblings names: Joe, Leah, and Anna Major

Personal medical history: No medical problems. Height - 5'6".

Family medical history: Her brother Joe has high blood pressure. Her father has back problems and her mother suffers from migraine headaches. Her sisters have no medical problems.

Genotype data:

1a	CGA	CAC	CAG	CAG	AAC	AGG	TCG	TGG	GCA	GCC
1b	CGA	CAC	CAG	CAG	AAC	AGG	TCG	TGG	GCA	GCC
2a	CCC	TGC	CGG	AGG	CTG	ACG	AAC	GTA	GTT	GCA
2b	CCC	TGC	CGG	AAG	CTG	ACG	AGC	GTA	GCT	GCA
3a	CAC	TGG	GTT	CCT	TGC	TAT	CGA	TCG	GTC	AGC
3b	CAC	TGG	GTT	CCT	TGC	TAT	CGA	TCG	GTC	AGC
4a	CAA	AGT	GAT	GTG	AAT	AGC	TTT	CAT	CTT	TAG
4b	CAA	AGT	GAT	GTG	AAT	AGC	TTT	CAT	CTT	TAG

APPENDIX L

Contract with Learning Systems Consultants, Inc.

AGREEMENT

THIS AGREEMENT, made and entered into this 17 day of August 1993, by and between the BIOLOGICAL SCIENCES CURRICULUM STUDY, hereinafter called BSCS, and LEARNING SYSTEMS CONSULTANTS, INC., hereinafter called LSCI, witnesseth:

1. Work to be completed by LSCI for THE HUMAN GENOME PROJECT: INFORMATION MANAGEMENT, ACCESS, AND REGULATION, hereinafter called HGN2, under this agreement includes:

a. LSCI will develop one (1) simulated database program ready for incorporation into the final edition of HGN2. The final delivery date for this simulated database is 1 May 1994. LSCI will develop the simulated database program based upon the prototype database program developed under the BSCS/LSCI Service Contract dated 10 July 1993, Attachment 1 (List of General Requirements), and Attachment 2 (List of Changes Resulting From the Writers' Conference) and subject to provisions in b - d below. In addition, LSCI will provide interim iterations of the product, as described in 1b, 1c, and 1d, below.

b. LSCI will provide a pilot product by 15 October 1993, for use by the BSCS staff in-house and at the NSTA regional workshop in Denver on 29 October 1993.

c. Based upon feedback from the use of the 15 October 1993 product, LSCI will revise the simulated database program by 15 November 1993, for use in the 2, 3

December 1993 teacher-orientation session and the January 1994 field test of the complete HGN2 instructional module.

d. Based on the results of external review and of the January 1994 field-test of the program revised per 1c, above, LSCI will revise the program for inclusion in the final distribution of HGN2. LSCI will deliver the revised program to BSCS by 1 May 1994. At a minimum, the revisions will encompass the following aspects of the program:

- Search on a FIELD or TYPE
- Compare function
- Password access
- Movement from screen to screen

2. In consideration of BSCS's obligations hereunder, LSCI:

a. Agrees to deliver to BSCS master materials and written descriptions for the field-test and final versions of the simulated database program, and documentation of the features of the program, as described in Attachment 3.

b. Agrees to complete all tasks related to the development of the simulated database program including, but not limited to, programming and debugging after data entry.

c. Agrees to revise the field-test version of the program based on feedback from BSCS staff, expert reviewers, and field-test data.

d. Warrants that it will obtain the permissions and rights to all extant materials it chooses for inclusion in the simulated database program independent of consultation with BSCS.

e. Grants to BSCS, its successors, and assigns exclusive rights to all new materials produced under this agreement.

f. Agrees to work with and consult the HGN2 staff and other specialists as BSCS shall direct, provided such consultation does not obligate LSCI beyond the scope of this agreement.

g. Agrees to help prepare, read, revise, correct, and return promptly any manuscript materials related to the use of the program.

h. Agrees to correct within 15 days any deficiencies found in the final product.

i. Agrees to deliver an MS-DOS version of the program by 1 July 1994.

j. Agrees to include the following statements on all versions of the program:

1. "This material is based on work supported by the United States Department of Energy under grant number DE-FG03-93ER61584. Any opinions, findings, and conclusions or recommendations expressed in the publication are those of the authors and do not necessarily reflect the views of the United States Department of

Energy."

2. "Copyright 1994 by BSCS. All rights reserved. You have the permission of BSCS to copy this software for your classroom use. The copyright on this software, however, does not cover reproduction of these items for any other use. For permissions and other rights under this copyright, please contact the Permissions Department, BSCS, 830 N. Tejon Street, Suite 405, Colorado Springs, Colorado 80903, U.S.A."

3. In consideration of LSCI's obligations hereunder, BSCS:

a. Agrees to develop student and teacher materials that include appropriate strategies for using the simulated database program as an integrated part of HGN2.

b. Agrees to enter all science-specific data for the HGN2 program.

c. Agrees to arrange for reviews and field tests as may be needed.

d. Agrees to provide consultation with the HGN2 staff and others as may be required.

e. Agrees to obtain all required permissions for all extant materials it includes in the database program.

f. Agrees to acknowledge LSCI's contributions on the database program itself and in the print materials that accompany the program. This acknowledgement shall read "Software development by Learning Systems Consultants, Inc., an educational technology

consulting firm and a developer of *learning systems for today's kids* -- Colorado Springs, CO, (719) 632-5450."

g. Agrees to pay LSCI four thousand dollars ($4,000.00) for the project as agreed to herein. See Attachment 4 for payment schedule.

h. Agrees to copyright all versions of the program in its own name.

All payments are contingent upon receipt of anticipated grant funds from the United States Department of Energy (DOE), upon timely performance by LSCI, and upon BSCS's approval of LSCI's work to date, which approval shall not be unreasonably withheld.

4. This work is developed under a DOE grant to BSCS, grant #DE-FG03-93ER61584, and all work under this agreement shall be subject to DOE guidelines, constraints, and limits.

a. Notwithstanding any other provisions of this agreement, failure of LSCI to perform, and to deliver required work, acceptable to BSCS, will result in the withholding of payment under this agreement unless such failure arises out of causes beyond the control, and without the fault or negligence of LSCI. BSCS shall promptly notify LSCI of its intention to withhold payment of any invoice or voucher submitted.

b. DOE, the Comptroller General of the United States, or any of their duly authorized representatives, shall have access to any books, documents, papers, and records of LSCI that are directly pertinent to this agreement for the purpose of making audits, examinations, excerpts and transcriptions.

c. Termination: By written notice at any time prior to BSCS's acceptance of the simulated database program, LSCI or BSCS may terminate this agreement. If either LSCI or BSCS terminates this agreement, LSCI shall transfer sole and complete rights to the terminated program to BSCS. All programming, visual, and script elements will become the sole property of BSCS. LSCI shall return to BSCS all payments previously made in payment therefor, excepting funds paid for documented costs associated with the project to that point. If BSCS terminates this agreement and does not plan to publish the terminated program, LSCI shall have the option, exercisable by written notice to BSCS within thirty (30) days after receipt of the BSCS termination notice, to purchase the terminated program upon payment to BSCS in cash within one hundred eighty (180) days of the LSCI notice of intent to purchase, a sum equal to all amounts paid by BSCS to LSCI for that program, including the amounts paid to LSCI pursuant to the BSCS/LSCI agreement of 10 July 1993. Should LSCI choose to publish said program, there shall be no reference to BSCS in the program or in associated marketing and promotion without the written consent of BSCS. Any publication of the terminated program by LSCI or its assignees shall result in payment of reasonable royalties to BSCS, at a rate to be negotiated by BSCS and LSCI prior to the time of publication.

5. This agreement shall be binding upon and inure to the benefit of the successors and assigns of LSCI and the successors and assigns of BSCS.

6. Nothing contained in this agreement shall be deemed to constitute the relationship between BSCS and LSCI as that of partners or joint venturers, nor principal and agent or employer and employee. BSCS and LSCI expressly agree their relationship is that of independent contractors.

IN WITNESS WHEREOF, the parties hereto have executed this Agreement the day and year first above written.

Biological Sciences Curriculum Study	Learning Systems Consultants, Inc.
BY: _____	BY: _____
Joseph D. McInerney, President	Jeff Thomas, President

ATTACHMENT 1
LIST OF GENERAL REQUIREMENTS

1. The database should have the following search capabilities:

   NGD: search by genetic trait, search by organism, search by genotype (label), search by sequence, search by chromosome number

   LGD: search by sample number. search by sex, search by sequence, search by genotype (label), search by personal data, search by family information

2. The database should have the following record types:

   NGD: may have multiple record types in which we will provide general information, allele information (genotype and phenotype), and map information

   LGD: individual records, family pedigrees

3. The database will have the following security features:

   LGD: individual records accessed only by password; there may be another set of passwords that give access to full family data; there may be a master password that gives access to all LGD data

4. Search results should display one record if the search argument is unique, and a list if the search argument is not unique. In the second case, the user should be able to select from the list to go to a unique record.

5. Additional design features:

   button that returns the user to the last record list

   compare functions: simple dot-plot and stacked

   ability to search the NGD directly from the LGD (sequence and genotype searches), and vice versa

ATTACHMENT 2
LIST OF CHANGES RESULTING FROM THE WRITERS' CONFERENCE

- In general, the screens should have the look that they now have. An entry for the LGD will look like the sample below. I'm not sure if we have as much room on the screen as needed. If necessary, we can shorten the Current status and possibly the Personal and Family medical histories.

Sample Number	37
Name:	Ann Thomas
Sex:	Female
Age:	41 (The above four lines could go on one line)
Current status:	Ann is. and always has been since her marriage to Harold, a homemaker. Twice a week she volunteers as a docent at the children's museum.
Parents names:	Norbert and Barbara Dudley
Siblings names:	Mary Beth, John, Peggy, Michael, Bridget, and James Dudley
Personal medical history:	Ann is a recovering alcoholic and spent 6 months in a psychiatric hospital soon after the birth of her first children when she was 23 years old. Height - 5'2".
Family medical history:	Ann's father suffered from alcoholism most of his life. Her mother died of a heart attack when Ann was 15. Except for Bridget, who is a recovering alcoholic, her siblings have no serious medical problems. Height 5'2".

Genotype data:

1m	CGA	CAC	CAG	CAG	AAC	AGG	TCG	TGG	GCA	GCC
1p	CGA	CAC	CAG	CAG	AAC	AGG	TCG	TGG	GCA	GCC
2m	CCC	TGC	CGG	AAG	CTG	ACG	AGC	GTA	GTT	GCA
2p	CCC	TGC	CGG	AAG	CTG	ACG	AGC	GTA	GCT	GCA
3m	CAC	TGG	GTT	CCT	TGC	TAT	CGA	TCG	GTC	AGC
3p	CAC	TGG	GTT	CCT	TGC	TAT	CGA	TCG	GTC	AGC
4m	CAA	AGT	GAT	GTG	AAT	AGC	TTT	CAT	CTT	TAG
4p	CAA	AGT	GAT	GTG	AAT	AGC	TTT	CAT	CTT	TAG

- The TYPE choice should have the following values. We might need to add others.

Database	TYPE	VALUE	View will give
LGD	Sample	1 to 52	Sample number, Sex, Age, Genotype
LGD	Sex	male and female	Sample number, age, Genotype
Both	Sequence	add either manually or from a click on a sequence in the	

		LGD
Both	Genotype	1a-1d 2a-2d 3a, 3b 4a, 4b or Enter Value
LGD	Personal Data	pop-up screen asking for name and password
NGD	Genetic Trait	Glaucoma Fragile x Height Hypertension or Enter Value
LGD	Family information:	Pedigree by name: Mota/El-Shanti/Chen/McCarthy Peters/Schmidt Thomas/Major/White Pedigree by genotype - two choices, with and without names
NGD	Organism	Enter Value, others to be determined

- The NGD has three sections, accessed by clicking on buttons in the choice bar.
  - GENERAL section has information about the trait. It fits on one screen.
  - ALLELE has the alleles, sequences, and phenotypes listed.
  - MAPPING data has one screen worth of information about the map location.

  If the student is in the GENERAL section, buttons in the choice bar should be ALLELE and MAPPING. If the student is in the ALLELE section, buttons should be GENERAL and MAPPING. If student is in the MAPPING section, buttons should be GENERAL and ALLELE.

- Need a CANCEL button for students who get into an area where access is limited.

- Need to have a button in the NGD that returns them to the previous LGD screen.

- Each database name will have a password assigned to it. When students access PERSONAL DATA through the TYPE choice they will see a screen that asks for Name and after the name is entered, the screen asks for the password. If incorrect information is entered, a pop-up message appears that states ERROR HAS OCCURRED and returns the menu to the original state. A global password of GENETIC COUNSELOR is needed for day two activities and day three activities. When GENETIC COUNSELOR is entered, the names on the database are shown.

- We need three pedigrees (one for each family) that are already part of the database. The pedigree will include several types of information, accessed at different times. The pedigree will have two forms. One will have the genotype, for each individual, without names but with sample numbers and the other will have genotype and the family and individual names. If students ask for the pedigree with names, the screen will ask for the password. The password is GENETIC COUNSELOR. One person in each extended family will have an incorrect genotype. An asterisk will be next to the name with the words INCOMPATIBLE GENOTYPE somewhere on the bottom of the screen.

- The search by sequence is maintained from the prototype, however, if the student double-clicks on the sequence in the LGD then a pop-up menu appears that asks if the student wants to search in the NGD by genotype. The search needs to be based on a "contains" algorithm. The view would look like:

Number	Genotype
2	2c
6	2c
13	2c
14	2c
16	2c
25	2c
45	2c

- The students will need to search the LGD for any sequence with the result that the number of the individual and the genotype is given. We do not need the sequence to appear on the screen after the search. (See above) If the password of GENETIC COUNSELOR is entered when the student logs into the LGD, the name will also appear. (See below)

Name	Sample Number	Genotype
Consuela Mota	2	4b
Anna McCarthy	8	4b
Jessie Peters	16	4b
Dan Peters	19	4b
Henry Thomas	32	4b
Laura Major	35	4b

- We will enter the data. If possible, a macro for doing so would be nice. All the LGD is now in Microsoft works database.

- We will have two different compare functions. One will have the simple dot-plot comparison. The other will have a stacked comparison, The compare function needs a different format than on the prototype. It should have:

  Standard sequence
  compared sequence 1
  Standard sequence (same as above)
  compared sequence 2
  etc.
  The standard sequence will always appear in its entirety, The compared sequence will only have those bases that are **different** from the standard appear. If possible, we'd like the numbers 1-30 to appear above the standard sequence in both types of compare. Font sizes might make this difficult to do.

```
1 2 3 4 5 6 7 8 9 10
A T G C C C T G A T
    A     T       G
```

ATTACHMENT 3

Required descriptions include all features of the program that would allow BSCS staff to write documentation for teachers and students. The descriptions include, but are not limited to:

- selection mechanisms such as choosing from buttons or pull-down screens;
- selection mechanisms such as using a mouse click for selecting data for searching in the LGD or NGD;
- method of going from one screen in the NGD to the next;
- method of selecting genotype information in the pedigree;
- any features of the simulation that make it user-friendly; and
- the means of going from the NGD to the previous LDG screen or vice versa.

The description of the database program may be in a narrative form but the features documentation shall specify each feature and means of using that feature. The main purpose of the features documentation is to avoid having undocumented features that might create problems or concerns for the student and teachers who use the simulation.

ATTACHMENT 4

The Human Genome Project:
Information Management, Access, and Regulation

Payment Schedule
for
Learning Systems Consultants, Inc.

Payment 1

upon signing of the contract	$1,500
upon receipt of pilot program (scheduled for 15 October 1993)	500
upon receipt of field-test version (scheduled for 15 November 1993)	1,000
upon receipt of final program (scheduled for 1 May 1994)	1,000
Total	$4,000

APPENDIX M

List of Potential Field-test Teachers

List of Prospective Field-test Teachers
for
HGN2

Anne L. Barnes
Tulsa Public School
Raymond S. McLain High School
4929 North Peoria Avenue
Tulsa, Oklahoma 74126

Gina Castro Brandt
P.O. Box 1532
Florence, Oregon 97439
(503) 997-9426

William Carbone
Glen Rock Junior-Senior High School
Glen Rock, New Jersey 07452

Alex Clark
Science Department Chairman
Fremont Ross High School
1100 North Street
Fremont, Ohio 43420
(419) 332-8221

Arline Deacon
511 Ginger Cake Road
Fayetteville, Georgia 30214
(404) 461-3576
school:
Woodward Academy
1662 Rugby Avenue
College Park, Georgia 30337
(404) 465-8217

Linda Dekort
Flathead High School
644 4th Avenue, West
Kalispell, Montana 59901
(406) 756-4502

Theresa Estes
Southwest Science/Math Magnet
6512 Wornall Rd.
Kansas City, Missouri 64113
(816) 871-0900

Aaron Feik
North Shore School District
18315 Bothell Way NE
Bothell, Washington 98011
(letter states they have numerous teachers who would like to field-test HGN2)

Sally K. House
Coconut Creek High School
The School Board of Broward County, Florida
1400 Northwest 44th Avenue
Coconut Creek, Florida 33066
(305) 977-2100

Nora K.W. Howard
Hilton Head High School
70 Wilborn Road
Hilton Head Island, South Carolina 29926
(800) 689-7550

Betty Hunt
Pius XI High School
135 N. 76th St.
Milwaukee, Wisconsin 53213
(414) 258-0532

Larry Jeffryes
103 Bandelier
Los Alamos, New Mexico 87544
(505) 672-9073
school:
Los Alamos High School
1300 Diamond Drive
Los Alamos, New Mexico 87544
(505) 662-4136

Dennis McKinney
Orcas Island High School
P.O. Box 167
Eastsound, Washington 98245
(206) 376-2287

Barbara Neureither
Holt High School
1784 Aurelius
Holt, Michigan 48842

Glenn Pickett
Science Department Chair
David H. Hickman High School
1104 North Providence Road
Columbia, Missouri 65203
(314) 886-2500

Gary Stellern
John Muir High School
1905 N. Lincoln
Pasadena, California 91104

Dr. Kaye Walter
Southwest Science/Math Magnet
6512 Wornall Rd.
Kansas City, Missouri 64113
(816) 871-0900

Jean K. Willard
270 Maple Street
Franklin, Massachusetts 02038

Dr. J. G. Yuhas
P.O. Box 222
Bar Mills, Maine 04004-0222

APPENDIX N

Cover Letter and Background Survey

Cover Letter for Field-Test Application

Thank you for your interest in field testing the new BSCS curriculum module *The Human Genome Project: Information Management, Access, and Regulation.* As Ed Drexler undoubtedly explained to you, the module includes background material for teachers on information technology as it relates to the Human Genome Project, and on ethical, legal, and social issues associated with the storage of genomic data in electronic databases. It also includes five days of classroom instruction that involve students directly in the manipulation of hypothetical research databases and registries, and in classroom analyses of ethical and public-policy questions related to the collection, dissemination, and use of genomic data.

At this time, we are looking for field-test teachers who are willing to use the materials in at least two introductory biology classes, with 20-30 students per class. Ideally, we would conduct the field test during January and February of 1994. Responsibilities of designated field-test teachers will be:

1. to attend a two-day orientation session at BSCS in Colorado Springs on 3-4 December 1993. BSCS will pay all expenses for travel, lodging, and meals, as well as a $250.00 per-day honorarium for the session. To contain budget costs and to have two full days available for discussion and interaction, we will ask teachers to arrive Thursday, 2 December, and to leave Sunday, 5 December.

2. to review the background information for teachers, and to teach all of the instructional activities included in the module. The first field-test version of the database is designed to run only on a Macintosh; as written, the activities can accommodate student/computer ratios from 30 to 1 to the recommended 4/1.

3. to collect detailed evaluation forms from each student on each activity.

4. to provide detailed feedback on each activity and the teacher background material, from the perspective of the teacher.

5. to allow the use of his or her name as well as the name of the school in the list of field-test sites that we will include in the final document.

6. to permit site visits by the project director, technology specialist, or project evaluator.

In addition, to meet our grant obligations, each field-test teacher must:

1. provide to the school board (or a comparable authority) information concerning

the need for and purposes of the project, the proposed content of the material, the expected benefits to be derived, and other information to assist the jurisdiction in arriving at a decision on participation.

2. obtain, after the responsible authority has carried out its procedures, written approval for participation in the project activities.

3. provide information and material to the responsible school authority to assist it in carrying out its own established procedures regarding the participation of students in project activities.

4. make available information and materials for inspection by parents or guardians of children engaged in the project.

Grant funds will allow us to offer support to only five designated field-test sites, though we have received expressions of interest from more than 20 schools nationally. If you would like to be considered as a potential field-test teacher, please complete the enclosed background survey and return it to BSCS by 1 October. Ed suggested that you might be interested in using the module with several of your sections of biology, and that you might be willing as well to test it under varying conditions (e.g., with different computer/student ratios). If this is the case, please indicate it on your survey. We will select sites that are broadly representative with respect to location, teachers, and students, and expect to be able to let you know the outcome of the selection process by 1 November 1993.

If you have questions about the project or the selection process, please do not hesitate to contact us. We appreciate your interest in BSCS, and trust that you will have a pleasant and productive school year.

Sincerely,

Joseph D. McInerney
Principal Investigator

Lynda B. Micikas
Project Director

JDM/LBM/dm
Encs.

BACKGROUND SURVEY
The Human Genome Project:
Information Management, Access, and Regulation

Name _____ Female ___ Male ___

Name of School _____

School Address _____

Home Address _____

School Telephone _____ Home Telephone _____

School FAX _____ Home FAX _____

Best time to call at school _____ at home _____

E-mail address _____

****

Professional Preparation and Experience:

Teaching experience: _____years teaching biology _____total years teaching

Type of degree: bachelor___ bachelor + hours___ master___
master + hours___ doctorate___

Have you received specialized training (preservice or inservice) in any of the following? (Please check all that apply.)

___genetics ___molecular biology ___genetic technology
___use of databases ___use of computers

What is your level of computer literacy? (List the *kinds of computers* and *programs*, [e.g., word processing, spreadsheets, databases] you are comfortable working with.)

_____
_____
_____
_____
_____

To what extent do you use computer-based technology in teaching? (List *frequency of use, ways in which you use such technologies, other.*)

Indicate which software packages you use personally and in the classroom.

(Personally)

(Classroom)

Have you taught the first BSCS genome module (*Mapping and Sequencing the Human Genome*)? ___yes ___no

If yes, what portion(s) of the module did you actually use?

What courses will you be teaching in the fall and winter of 1993-1994?

What biology textbook(s) do you currently use?

Computer Support:

Can you communicate with an electronic bulletin board? ____

Can you make long distance calls from your school if the school is reimbursed at a later date? _____

Is there a phone line available near your computer to which a modem can be connected? _____

What kinds of computer equipment (e.g., *computers*, *modem*, *printer*) are available to you at school for downloading?

_____

_____

_____

What kinds of computer equipment are available to you at school for instruction? (List how many of each are available, and whether they are available for use in the classroom or in a computer lab.)

_____

_____

_____

_____

What would be the ratio of students/computers if you were to teach this module?

_____

School Characteristics:

Check all of the following that apply to your school:

___urban ___suburban ___rural ___public ___private

Total enrollment in your school: _____ Percentage of minority students: _____

Which ethnic group(s) do these minority students represent? _____

With how many class sections would you like to field test this module? _____

How many total students would be involved in the field test? _____

APPENDIX O

Preliminary Agenda for the Field-test Orientation

AGENDA

The Human Genome Project:
Information Management, Access, and Regulation

Field-Test Orientation Meeting
3,4 December 1993
Antlers Doubletree Hotel
Colorado Springs, Colorado

Friday, 3 December

8:00 - 8:30 a.m.	Continental breakfast
8:30 - 8:45 a.m.	Welcome, introductions, and review of agenda (J. McInerney)
8:45 - 9:00 a.m.	Overview of project and time line (J.McInerney)
9:00 - 9:30 a.m.	General overview of experimental materials (L. Micikas)
9:30 - 9:45 a.m.	Questions and discussion
9:45 - 10:00 a.m.	Break
10:00 - 10:30 a.m.	Review of teacher narrative (J. McInerney)
10:30 - 11:30 a.m.	Demonstration of model databases (P. Goulding)
11:30 a.m. - 12:45 p.m.	Lunch
12:45 - 1:45 p.m.	Activity One (J. Murray, E. Drexler)
1:45 - 2:45 p.m.	Activity Two (J. Murray, E. Drexler)
2:45 - 3:15 p.m.	Break
3:15 - 4:15 p.m.	Activity Three
4:15 - 4:45 p.m.	General questions and discussion
5:00 p.m.	Reception
6:00 p.m.	Dinner

Saturday, 4 December

8:00 - 8:30 a.m.	Continental breakfast
8:30 - 9:30 a.m.	Activity Four (M. Cutter, L. McCullough, J. Zola)
9:30 - 9:45 a.m.	Break
9:45 - 10:45 a.m.	Activity Five (M. Cutter, L. McCullough, J. Zola)
10:45 - 11:30 a.m.	General questions and discussion
11:30 a.m. - 12:45 p.m.	Lunch
12:45 - 2:00 p.m.	The field test: evaluation of teacher and student materials (R. Backe)
2:00 - 2:30 p.m.	Break
2:30 - 3:30 p.m.	The field test: logistics (L. Micikas) - print materials - database - returning evaluation forms - telephone interviews - help with problems - site visit by project staff
3:30 - 4:00 p.m.	General questions and discussion
	Adjourn

APPENDIX P

Letter to Daniel W. Drell, Department of Energy, Washington, DC

BSCS

Innovative Science Education

founded 1958

5 January 1993

Daniel W. Drell, Ph.D.
Health Effects and Life Sciences
Research Division, ER-72
Office of Health and Environmental
Research, Office of Energy Research
Department of Energy
Washington, DC 20585

Dear Dan:

I hope this finds you well and rested following the holidays and looking forward to an interesting and productive 1993. Certainly, things will be interesting for you and your colleagues at DOE as you accommodate to the new administration and a new Secretary of Energy. One wonders whether Hazel O'Leary will have the same level of commitment to science education as did Admiral Watkins.

Enclosed, per our recent telephone conversations, is a revised budget for our proposed follow-up to the BSCS module on the Human Genome Project. The reduction is $104,251 from the original (22.3%). I am in a bit of a bind -- pedagogically as well as financially. The original budget does not contain much fat. I realize you must hear that all the time, but -- I feel like Dave Barry -- I am not making this up. I have restructured the funding periods so that year 1 is now 6 months (1 April - 30 September 1993, timed to end with the conclusion of the government's fiscal year) and year 2 is 16 months (1 October 1993 - 31 January 1995).

I have reduced the budget by omitting the costs of printing and distributing the print materials intended to accompany the sequence data bases that we will distribute electronically (over the BSCS bulletin board). We can make the print materials available through the same mechanism. This option, however, exacerbates some concerns about pedagogy and access, most notably the concern of your reviewers that the project may be limited to elite schools by virtue of its reliance on educational technology and telecommunications. In addition, by providing the print materials on the bulletin board we constrain severely the use of graphics because schools may not have printers that accommodate the types of graphics we hope to use to illustrate the structure of sequence data bases and other issues related to the science and technology of the HGP. Last, if we load the print martials on the bulletin board I cannot reduce the FTE for our technical specialist because he (Phil Goulding) will be critical for maintaining the large amount of material

on the board and for answering the inevitably large number of questions that teachers will have as they try to download the expanded volume of material.

Unless I find outside (private sector) support for the print materials I can think of no other way to reduce the budget significantly without constraining the original intent of the project. I will welcome your feedback on this matter.

You shared with me some other concerns the reviewers expressed about the proposal, and I shall address those below.

Concern 1. The reviewers could not fully evaluate the proposed conceptual framework for the program.

Response: The content proposed on pages 6-10 of the proposal is provided simply to indicate issues that we at BSCS feel are important and that might be included in the final product. (Obviously, these topics also are listed to show the reviewers that we are conversant with the major issues.) The BSCS process depends heavily upon input from the advisory committee to help determine the conceptual framework. That is why we choose the committee members so carefully and pay such close attention to their deliberations. Although I have a general sense of what I hope to accomplish, I will take substantial guidance from the committee in developing the conceptual framework that the writers then will turn into classroom instruction.

Concern 2. The writers are listed only as <u>potential</u> writers.

Response: It is difficult to get busy people to commit to two hypothetical weeks. I have, however, spoken with all of the people listed in section 3.3 and all are extremely interested in coming to Colorado Springs to work with us. Most of them, in fact, worked on the first module and are enthusiastic about the chance to work together again and to contribute to public understanding of the HGP.

Concern 3. The procedure for selecting test sites in unclear.

Response: Section 3.6 of the proposal explains the process only briefly, but appendix E includes the survey questions we use to help us select the sites. We try to have geographic and ethnic/racial balance, and we try to ensure that we have urban, suburban, rural, public, and private schools, as well. It is difficult, however, to cover all of those variables when the cost of field-testing drives the budget up so quickly. As the proposal states, we get many more requests for field testing than the budget possibly can accommodate, so we select on the basis of the foregoing data and with the intent of having a few experienced BSCS sites and a few new ones.

Concern 4. It is not clear that we will address ELSI issues sufficiently.

Response: Sections 2.2.2 and 2.2.3 of the proposal (pages 7 and 8) list some of the issues we hope to address. Perhaps the reviewers were influenced by the amount of time the proposal devotes to the science and technology of the relevant data bases. BSCS long has been committed to introducing teachers and students to ELSI-type issues, but we also believe that a sound understanding of the relevant science must underlie any analyses of ethics and policy. We always struggle with the appropriate balance, and I think we achieved a good mix in the current module.

Concern 5. The use of educational technology and telecommunications may limit this project to use by elite schools.

Response: Pages 9 and 10 of the proposal provide some data on the use of computers in American high schools. The number of computers is large and continues to grow. In addition, we feel that the use of the free 800 number to access the BSCS bulletin board will overcome many of the barriers related to the cost of long-distance phone calls and the need for dedicated lines. Were we to distribute the sequence data on discs in the three required formats, the budget would increase by about $40,000. I realize that the approach we have proposed has some trade-offs, but we believe they are justified.

Please let me know what you think about my budget dilemma. In addition, please contact me should you require more detail than I have provided in my response to the foregoing concerns expressed by the reviewers. Thanks for your continued help and support. Best regards.

Sincerely,

Joseph D. McInerney
Director

JDM/dm
cc: K. Winternitz
L. Satkowiak

DATE
FILMED
5/3/94
END