NATIONAL ADVISORY COUNCIL FOR HUMAN GENOME RESEARCH
SUMMARY OF MEETING[1]
September 11, 2006

The open session of the National Advisory Council for Human Genome Research was convened for its thirty-eighth meeting at 8:34 A.M. on September 11, 2006 at the Fishers Lane Conference Center, Rockville, MD. Francis Collins, Director of the National Human Genome Research Institute, called the meeting to order.

The meeting was open to the public from 8:34 A.M. until 12:45 P.M. on September 11, 2006. In accordance with the provisions of Public Law 92-463, the meeting was closed to the public from 12:45 P.M. on September 11, 2006 until adjournment for the review, discussion, and evaluation of grant applications.

Council members present:

Andrew Clark
Marilyn Coors
Geoffrey Duyk
Sean Eddy
Vanessa Gamble
William Gelbart, *ad hoc*
Deidre Meldrum, by teleconference
Jeffrey Murray
Thomas Murray, *ad hoc*
Stephen Prescott
Harold Shapiro
George Weinstock

Council members absent:

Beverly Gaines
Mary Hendrix

Ex Officio member absent:

Gerard Schellenberg

Staff from the National Human Genome Research Institute:

[1] For the record, it is noted that to avoid a conflict of interest, Council members absent themselves from the meeting when the Council discusses applications from their respective institutions or in which a conflict of interest may occur. Members are asked to sign a statement to this effect. This does not apply to "en bloc".

Solome Abebe, DER
Catherine Bennet, DER
Saveri Bhattacharya, DER
Christianne Bird, DER
Vivien Bonazzi, DER
Joy Boyer, DER
Lisa Brooks, DER
Amanda Broadnax, DER
Comfort Browne, DER
Cheryl Chick, DER
Monika Christman, DER
Francis Collins, OD
Chris Davis, OD
Karen DeLeon, OD
Gwendolyn Dudley, DER
Adam Felsenfeld, DER
Colin Fletcher, DER
Phyllis Frosst, OD
Peter Good, DER
Bettie Graham, DER
Alan Guttmacher, OD
Mark Guyer, DER
Emily Harris, DER
Laura Liefer, DER
Carson Loomis, DER
Teri Manolio, DER
Jean McEwen, DER
Keith McKenney, DER
James McWilliams, DER
Jessica Melone, DER
Ken Nakamura, DER
Kenneth Ow, OD
Brad Ozenberger, DER
Jane Peterson, DER
Rudy Pozzatti, DER
Michael Rackover, OD
Ed Ramos, OD
Eddie Rivera, OD
Cristen Robinson, DER
Anna Rossoshek, DER
Jeff Schloss, DER
Geoff Spencer, OD
Shundel Stephenson, DER
Tanya Stevens, OD
Gary Temple, DER
Elizabeth Thomson, DER
Fred Walker, OD
Kris Wetterstrand, DER

Others present for all or a portion of the meeting:

Judith Benkendorf, American College of Medical Genetics
Joann Boughman, American Society of Human Genetics
Susan Castillo, SRA International
Sharon Olsen, International Society of Nurses in Genetics
Carmen Perera, DEAS
Sharon Terry, Genetic Alliance
Wendy Uhlmann, National Society of Genetic Counselors
Diane Williams-Bey, DEAS

INTRODUCTION OF NEW MEMBERS AND STAFF, LIAISONS AND GUESTS

Dr. Guyer introduced a new council member: Harold Shapiro from Princeton University.

Dr. Guyer introduced two ad hoc council members: Thomas Murray from Hastings Center and William Gelbart from Harvard University.

Dr. Guyer introduced new NHGRI staff: Vivien Bonazzi, Program Director for Bioinformatics; Emily Harris, Program Director for Population Genomics; Ajay, Program

Director for Cheminformatics; Solome Abebe, Catherine Bennet and Cristen Robinson, Program Analysts; Anna Rossoshek, Scientific Administrator; and Ed Ramos, ASHG/NHGRI Policy Fellow.

Dr. Guyer welcomed members of the press and liaisons from professional societies: Joann Boughman from the American Society of Human Genetics, Sharon Terry from the Genetic Alliance, Wendy Uhlmann from the National Society of Genetic Counselors, Sharon Olsen from the International Society of Nurses in Genetics and Judith Benkendorf from the American College of Medical Genetics.

APPROVAL OF MINUTES

The minutes from the May 2006 Council meeting were approved as submitted.

FUTURE MEETING DATES

The following dates were proposed for future meetings: February 12-13, 2007, May 21-22, 2007, September 10-11, 2007, February 11-12, 2008, May 19-20, 2008 and September 8-9, 2008.

DIRECTOR'S REPORT

I. GENERAL ANNOUNCEMENTS

Dr. Collins noted the anniversary of the September 11, 2001 terrorist attacks on the United States and offered a few reflections. He remembered that the Council was in session when word of the attacks came. Most of the Council members who were present that day have since rotated off, but those that were involved in that Council meeting will always have that connection. Dr. Collins remarked that September 11th was an event that changed our perspective as a nation and rendered us a different people. We have been changed by what happened, and are a little less naïve about our own fragility. We should remember the more than 3,000 people who died in the World Trade Center, and the nearly 200 people who lost their lives in the crash at the Pentagon, and their families. The NHGRI had a direct connection with subsequent events, as several NHGRI staff members assisted in the identification of the remains of the victims at the World Trade Center.

A moment of silence was held at 8:46 am, the time when the first plane struck the North Tower of the World Trade Center.

In news, former council member Dr. Wylie Burke has been appointed as President-Elect of the American Society of Human Genetics (ASHG). Dr. Burke is Professor and Chair of the Department of Medical History and Ethics at the University of Washington in Seattle WA, and is the Principal Investigator of the Center for Genomics and Healthcare Equality, one of the NHGRI Centers of Excellence in ELSI Research (CEERs).

Another former Council member, Kim Nickerson, Ph.D., has joined the University of Maryland as an Assistant Dean in the College of Behavioral and Social Science (BSOS) and Director of the UMD BSOS Diversity Initiative. Dr. Nickerson continues to be involved with the NHGRI as one of the Institute's research training advisors.

Vice Admiral Richard H. Carmona's tenure as the 17th Surgeon General of the United States Public Health Service ended in early August, and he is returning to civilian life. Dr. Carmona was a strong supporter of the Family History Initiative, and has already met with Alan Guttmacher to express his interest in continuing work on this initiative. Dr. Carmona is succeeded by Rear Admiral Kenneth P. Moritsugu, MD, MPH, who has been Deputy Surgeon General of the United States since October 1, 1998.

NHGRI has begun recruitment for a new executive officer, after the Institute's current EO, Mr. Fred Walker, announced that he will be retiring in the winter of 2007. Dr. Collins expressed his appreciation for Mr. Walker's service, noting that he has been a wonderful contributor to the leadership of the institute, and has managed many challenging projects during his tenure.

II. NEW NHGRI INITIATIVES

A Request for Applications (RFA) for Technology Development for the ENCODE project has been reissued. Technology development is a continuing component of the ENCODE (Encyclopedia of DNA Elements) Project, which has the long-term goal of identifying all of the functional elements in the human genome. The purpose of the new RFA is to solicit another set of proposals to develop new and improved technologies for the efficient, comprehensive, and high-throughput identification and validation of all types of sequence-based functional elements in eukaryotic genomes. Areas of interest include projects to develop technologies to identify new types of functional elements, to identify functional elements in repetitive sequences, to validate functional elements that are being identified using existing methods, and that can be applied using small sample sizes. Both experimental and computational methods are being encouraged. RFA-HG-07-028 will use the R21 (Exploratory/Developmental) grant mechanism, and RFA-HG-07-028 will use the R01 grant mechanism. The receipt date for letters of intent receipt is October 30, and the application receipt date is November 28, 2006.

III. EXTRAMURAL PROGRAM

Sequencing Progress: Council was provided with a summary of the current status of GenBank submissions from the large-scale sequencing program. All of the projects currently being worked on were approved by Council.

New Sequencing Targets: A press release announcing NHGRI's latest sequencing targets, which were approved at the May 2006 Council meeting, was issued on July 19, 2006. The projects include full shotgun coverage of the gibbon genome, increasing the coverage of the genomes of seven mammals (armadillo, cat, guinea pig, elephant, little brown bat, tree shrew and rabbit) from 2- to 6-fold, and full shotgun coverage of the genomes of five Dermatophyte fungi that commonly infect humans. Two other projects that were approved were one to gather data to study population genomics by sequencing a large number of *S. cerevisiae strains* (full

shotgun of 10 and two-fold shotgun of 25; this project is being done in collaboration with the Sanger Institute) and one designed to provide data to study the origins of multicellularity (full shotgun of 6 species, low coverage of 4 species; includes fungi and protists).

Completed Projects: The genome sequencing groups and the organism communities for both the honey bee and the sea urchin are working on publications describing the sequencing and first analyses of these genomes. The Honey Bee Genome Consortium's analysis of the honey bee genome has been accepted by *Nature* and should be published at the end of October. An additional thirty companion papers on the honey bee genome have been accepted to other journals, including *Science, Genome Research,* and *Insect Molecular Biology.* Similarly, the sea urchin genome sequence paper is also expected to be accompanied by thirty analysis papers on the sea urchin genome Dr. Collins offered congratulations to the Baylor Genome Sequencing Center, which led both the honey bee and sea urchin efforts.

Maintaining the Human Genome Sequence: As agreed upon at a meeting held at Cold Spring Harbor in May 2006, the Human Genome Reference Committee has been established to coordinate efforts for continued curation of the human genome sequence. Curation is needed because the current version of the human sequence still contains recalcitrant gaps, which groups are still working to close, and may contain errors that need to be corrected. The committee membership consists of representatives from the Baylor College of Medicine, The Broad Institute, Genoscope, Riken, the Sanger Institute, Washington University, the European Bioinformatics Institute (EBI), the National Center for Biotechnology Information (NCBI), the University of California, Santa Cruz (UCSC) and NHGRI.

Centers for Excellence in Genomic Science: On August 22nd, 2006, the NHGRI announced the awarding of grants totaling $54 million over five years to establish one new Center of Excellence in Genomic Science (CEGS) and to renew support for two of the existing Centers. The CEGS program, which was started in 2001, supports interdisciplinary teams of scientists working on projects directed toward making critical advances in genomic research. The first centers were funded as five-year awards that are scheduled to end this fall. After competitive review, the NHGRI decided to renew the awards for the Microscale Life Sciences Center at the University of Washington, Seattle (Deirdre R. Meldrum, P.I.) and the Yale Center of Excellence in Genomic Science (Michael P. Snyder, P.I.). In addition, NHGRI awarded a new award for a CEGS at the California Institute of Technology, Pasadena, Calif. (Marianne Bronner-Fraser, P.I.), which will be called the Center for In Toto Genomic Analysis of Vertebrate Development.

The Cancer Genome Atlas (TCGA): This exciting partnership between NHGRI and NCI is about to launch its pilot phase. TCGA is an effort to apply a variety of genomic tools at scale to obtain a comprehensive description of all genomic changes in all major cancers. It will be a complicated project with many challenges, ranging from new consent issues, obtaining well-characterized tumor samples to the heterogeneity of most cancer tissues. NHGRI and NCI have organized an External Scientific Committee, co-chaired by Dr. Ronald DePinho of the Dana-Farber Cancer Institute and Council member Geoff Duyk, to provide on-going evaluation of the project's progress.

The process for selecting the tumor types for the pilot phase of TCGA has been underway over the past few months and NCI will make an announcement later this week of three types that will be studied initially. Over the past summer, proposals and applications for projects to participate in the TCGA were received and reviewed, and announcements of awards will be made in the next few months. The components that were competed include the Biospecimen Core Resource (BCR), which will take in the selected tumor types and distribute materials to the other parts of the project, and will be funded as a contract; the Cancer Genome Characterization Centers (CGCCs), which will perform various whole-genome analyses on the tumor samples to identify regions that are altered in the tumors, compared to normal samples; the NHGRI-supported Sequencing Centers that will perform PCR-directed sequencing in selected genes or target regions identified by the CGCCs: and a Data Coordination Center (DCC) that will track and distribute data from all production centers. The NHGRI contribution to the TCGA will consist of $50M of sequencing capacity over the next three years.

As a preliminary study, the NHGRI allowed the current sequencing centers to use existing funds to undertake technical demonstration projects to help develop the directed sequencing pipelines that will be needed for the TCGA.The Tumor Sequencing Project (TSP) Consortium is a collaboration among groups at the Baylor College of Medicine Human Genome Sequencing Center, the Broad Institute Genome Sequencing Platform, the Dana Farber Cancer Institute, the Memorial Sloan-Kettering Cancer Center, the Genome Sequencing Center and Siteman Cancer Center at Washington University, the M.D. Anderson Cancer Center and the University of Michigan Medical Center. The TSP will pilot approaches to large-scale identification of genomic changes in tumors, aiming to sequence the exonic regions of 1,000 genes in almost 200 specimens of adenocarcinoma of the lung, as well as to use high density SNP genotyping arrays for high resolution identification of changes in chromosomal copy number. A second collaboration, between investigators at the J. Craig Venter Institute (JCVI) and The Johns Hopkins University, will evaluate different technologies for sequencing tumor DNA. This project will analyze the DNA sequence of 37 genes in a collection of 20 glioblastoma tumors. More details about both of these demonstration projects can be found at www.genome.gov/cancersequencing .

NIH Roadmap: Participants in two Roadmap initiatives in chemical genomics, the Molecular Libraries Screening Center Network (MLSCN) and the Exploratory Centers for Cheminformatics Research (ECCR) met together during the annual MLSCN meeting in Washington, D.C. on July 17-19, 2006. The MLSCN centers, which have been functioning for a little more than one year at this point, have focused on their goal of providing small molecule technology to academia, developing a small molecule resource and developing better assays. The Network includes 10 screening centers , one of which is the intramural NIH Chemical Genomics Center (NCGC). To date, 135 assay proposals have been submitted to the MLSCN, 74 (57%) of which have been accepted after peer review. Forty-five of the assays have reached the primary screening stage and the screening data have been entered into PubChem. Of the 45, nineteen were done at the NCGC. A mid-course evaluation of the Molecular Libraries Initiative will be conducted in late 2006, as part of the decision-making for transition of the MLSCN program to full-scale operation. This evaluation will serve as a pilot for an assessment process that can then be used for all of the components of the Roadmap at such transition points.

On July 24th, a team from the NCGC published a paper in *Proceedings of the National Academy of Sciences* describing a new approach to high-throughput screening. Traditional high-throughput screening initially measures the biological activity of chemical compounds at a single concentration. The new approach, called quantitative high-throughput screening, or qHTS, tests the compounds at seven or more concentration levels spanning four orders of magnitude. qHTS offers a significant increase in screening efficiency and accuracy. This publication was the most highly downloaded PNAS paper for the month of August.

ENCODE: The ENCODE pilot project is about to complete its third year. The scale up of ENCODE was postponed to extend the pilot for a fourth year according to the recommendations of a meeting of the ENCODE advisors that was held in January 2006 and were then approved by Council. An ENCODE Consortium meeting in July brought together Consortium members, the ENCODE Scientific Advisory Panel, and other investigators who are working on ENCODE-related projects. The meeting focused on the scientific aspects of the Project, and on issues related to measuring data quality, scaling the project to the entire human genome, and defining an endpoint for the project. Specific attention was paid to one of the unique challenges being addressed in the ENCODE pilot, managing, coordinating and analyzing many different data types. A publication describing the scientific accomplishments of the ENCODE pilot project to date is under development.

Applications for the model organism ENCODE (modENCODE) Project, which will seek to identify elements in the *C. elegans* and/or *Drosophila melanogaster* genomes, have been received in response to RFAs HG-06-006 (modENCODE projects) and HG-06-007 (a modENCODE Data Coordination Center). The applications will be reviewed in November and brought to February 2007 Council for funding in early spring 2007.

Knockout Mouse Project (KOMP): A June press release announced the launch of the NIH effort to construct a knockout mutation in every gene in the mouse genome and make the library available in public repositories. This project is jointly funded by many institutes, but is not part of the NIH Roadmap. The specific goal of the KOMP is to build a comprehensive and publicly available resource of knockout mutations for each of approximately 20,000 protein-coding genes in the mouse genome. Awards for projects to construct the mutations were made to Regeneron Pharmaceuticals, Inc. (David Valenzuela, P.I.) and a collaborative team from Children's Hospital Oakland Research Institute (CHORI), University of California, Davis (UC Davis); and the Wellcome Trust Sanger Institute (Pieter de Jong, P.I.). The CHORI-led team plans to create mouse embryonic stem (ES) cell lines in which 5,000 genes have been knocked out by gene targeting. Regeneron will aim at a different set of 3,500 genes. The two groups will use different knockout strategies for the first two years of the project, after which a decision will be made about the optimal approach to finishing the task.

In addition to the production effort, there are three other components of KOMP. Awards were made by the National Institute on Drug Abuse (NIDA) to the University of Pennsylvania (Klaus Kaestner, P.I.) and the Samuel Lunenfeld Research Institute (Andras Nagy, P.I.) to develop methods to create ES cell lines derived from the C57BL/6 strain that are suitable for

high-throughput gene targeting . In addition, the Regeneron award includes funds to optimize its existing C57BL/6 ES cell line. Second, The Jackson Laboratory was funded to establish a Data Coordination Center (Martin Ringwald, P.I.) that will collect, display and distribute information that will allow the research community to track the progress of knockout production. Finally, an RFA for the last component of KOMP, a repository, will be released shortly for funding in FY2007.

KOMP will closely coordinate its activities with two other large-scale mouse knockout efforts now underway, the North American Conditional Mouse Mutagenesis Project (NorCOMM; Canada), and European Conditional Mouse Mutagenesis Program (EUCOMM; Europe).

Mammalian Gene Collection (MGC): The total non-redundant MGC now consists of a set of full open reading frame clones corresponding to 14,162 human, 13,304 mouse, and 4,773 rat genes. MGC will next seek to use DNA synthesis to obtain clones for the approximately 2,300 total human and mouse genes that have eluded PCR rescue. An RFP for the DNA synthesis project will be issued in the coming weeks. A number of other possibilities for the use of the remaining MGC funds, including starting on the collection of alternate splice forms and converting existing ORFs to an expression-ready format are at different stages of planning.

Other Developments: A paper published in Nature last month by NHGRI-supported investigators provided an example of how genome sequences of organisms can be used for further research. David Haussler and his co-authors reported evidence for the involvement of a key gene, termed HAR1F (human accelerated region), in the evolution of the human brain. HAR1F is an RNA gene (non-coding) expressed in the developing human brain. In its sequence, 18 of the 118 nucleotides have changed since the human lineage separated from that of the chimp, which is many more changes than would be expected by neutral theory. The authors termed regions of the genome displaying this phenomenon "human accelerated regions" and hypothesized that such regions are related to human-specific biology.

IV. INTRAMURAL PROGRAM

On June 7th, NIH announced that nearly $4 million had been awarded to fund 19 bench-to-bedside medical research projects designed to speed translation of promising laboratory discoveries into new medical treatments. For the first time, applications for these awards, first given in 1999, were open to research teams made up of NIH intramural and extramural collaborators from medical schools, health-care organizations and private industry. Extramural scientists will take advantage of the NIH Clinical Center.

A paper in the *Journal of Cancer Research* described the results of a large study that has provided the clearest picture yet of the prevalence in the U.S. population of mutations in two genes associated with an increased risk of breast cancer, BRCA1 and 2. Elaine Ostrander,

Chief of the Cancer Genetics Branch in NHGRI's Division of Intramural Research, is one of the lead authors on the study.

NHGRI researchers have found that mutations in the gene for glucocerebrosidase may, in addition to causing Gaucher's Disease, also be an important risk factor for the second most common form of dementia among the elderly, dementia with Lewy bodies. The work was led by Dr. Ellen Sidransky, acting chief of the Medical Genetics Branch, NHGRI.

Drs. Robert Nussbaum and Jennifer Puck recently left the intramural program to take positions at the University of California, San Francisco. After a national search, Dr. David Bodine has been appointed as the new Chief of the Genetics and Molecular Biology Branch, and Dr. Leslie Biesecker has been appointed as the new Chief of the Genetic Disease Research Branch.

There are planned recruitment efforts in 2006-2007 for as many as 4 tenure-track investigators.

OFFICE OF THE DIRECTOR

GWAS Policy: NIH has issued a Notice for Public Comment, seeking comment on a proposed policy on genome-wide association studies (GWAS). Stimulated by the availability of the HapMap, many such studies have been initiated or are currently being planned. In an attempt to ensure as much consistency among them as possible and to potentiate the ability to do meta analyses across them, Dr. Zerhouni has called for the development of an NIH-wide policy for data sharing in such studies. A trans-NIH working group, led by NHLBI Director Betsy Nabel, has developed a draft policy for GWAS that addresses both the goal of rapid data access by investigators, which is required to advance this field rapidly, and the goal of maintaining the highest level of patient protection. The proposed NIH policy, especially the aspects regarding data access, is modeled closely on the policy developed for GAIN (see below). Policies developed for NIH programs such as the Genes and Environment Initiative (GEI) and the NHGRI's medical sequencing effort, will need to be consistent with the final NIH GWAS policy

As drafted, the GWAS policies will have many implications for investigators and the general public, and the notice seeks feedback from the general public as well as the scientific community. Comments on the proposed policy are due by October 31, 2006. Dr. Collins encouraged Council members to respond to the Notice with their thoughts. NIH hopes to finalize the policy by March 2007.

Council members raised questions about how to broaden efforts for communicating with the public and obtaining feedback on proposed policies. They noted that these policies will reach far outside of the scientific community, which means that it is very important that they are understood widely. When policy discussions are covered by the press, there is often a degree of misinformation provided, especially when the guidelines are complex and potentially controversial. If they are not well understood, there can be a delay in determining the true meaning of policy proposals and the adoption and implementation of

reasonable policies. It was suggested that NIH needs a strong effort to provide education to the research community when the guidelines are established.

Dr. Collins responded by noting that the NIH communication offices have discussed how to proactively communicate with the public, so misunderstandings are limited. For example, the ASHG will also discuss this issue during its Board of Directors meeting in October 2006, and the NIH will hold a town meeting for public discussion of the issues.

Council also noted that the possibility of placing these kinds of genetic data in the public domain raises many issues. As more studies are conducted, more reference sets will become available, increasing the opportunity for efforts to match up results from different studies. If genetic variation is matched to phenotype, these data could be used to make predictions about individuals. Dr. Collins replied that the proposed NIH policy attempts to address the data access issues and the potential misuse of data.

Identifiability Workshop: In a related matter, on October 3rd -4th, 2006, NHGRI will convene a small workshop to discuss the issue of identifiability in genomic research. The workshop is being organized by Dr. William Lowrance, an expert in confidentiality issues who is currently consulting with NHGRI. The workshop will consider such issues as the risks associated with identification, matching and probabilistic profiling as modes for identifying individuals from genotype data, technical options for de-identifying data, the implications of the Common Rule and the HIPAA Privacy Rule, and open publication of genomic data as compared with controlled release.

Short Course: The annual Current Topics in Genomic Research Short Course was held at NHGRI July 30-August 4, 2006. Seventeen faculty members and sixteen students from minority-serving institutions attended. Participants spent one week at NHGRI hearing talks from NHGRI faculty, touring labs, and meeting one-on-one with researchers. This program includes opportunities for NHGRI to maintain contact with the participating individuals and their institutions and to help them define their genomic curricula.

Community Genetics Forum: On September 14-16, the annual Community Genetics Forum will be held in North Carolina. This event had previously been held in Washington, D.C., but in 2004 NHGRI decided to convene future meetings outside of the D.C. area. The 2005 meeting was held in Seattle, WA. The 2006 location was chosen by a competitive process. The successful proposal was presented by a combination of investigators at University of North Carolina, Chapel Hill and Duke University. Sixteen NHGRI staff will travel to North Carolina to lead various sessions and forum activities.

NHGRI POLICY

Appropriations: The House Labor-HHS Appropriations Subcommittee took up the Labor-HHS appropriations bill in June, which includes a proposed 2007 funding level of $482,942,000 for NHGRI, which is approximately $3 million less than the 2006 budget. The only appropriations bill not passed by the House by the summer recess was the Labor-HHS bill (it was postponed after House Democrats attached minimum wage language in an effort to force a vote on that

issue). The House bill discusses TCGA, GAIN and GEI, and encourages follow-up on the AGES working group.

The Senate bill was passed in July. In it, NHGRI received a modest increase from the House level, to $486,315,000. The Senate bill also includes directive language and asks us to focus on liver disease, Parkinson's disease, SMA, and Tuberous Sclerosis Complex.

GOP leaders have vowed to clamp down on domestic spending, but Democrats and moderates oppose cuts to health and education programs funded by the Labor-HHS bill. For the first time in three years, no spending bills were ready for the President to sign before the August recess. An appropriation may not get approved until after the election. The proposed $40M for GEI is expected to survive the process.

Genetic Non-Discrimination: As of the August recess, there were 230 co-sponsors of the Genetic Nondiscrimination Information Act, H.R. 1227. The Coalition for Genetic Fairness, led by Sharon Terry of the Genetic Alliance, has been working with Congresswoman Judy Biggert's staff in an effort to move the bill through the House committees and towards a vote this session. The Coalition has made steady progress on employment provisions, but there is very little time left to get the bill finished before the elections.

SACGHS: The Secretary's Advisory Committee on Genetics, Health and Society met for its summer meeting on June 26-27. The committee's draft report on Large Population studies and the policy issues involved was released for public comment in late May. Since then they have received many comments, many of which address the importance of a national study. The committee discussed policy issues relating to assessment of environmental components of gene-environment studies, the NAS intellectual property report, direct-to-consumer marketing of genetic test kits and the FDA-FTC advisory warning.

Judy Yost of CMS gave a presentation to the SACGHS about the Notice of Proposed Rule Making (NPRM) related to the 2001 Notice of Intent to promulgate a rule relating to a genetic specialty area under CLIA, and advised the committee that a draft rule was in the clearance process at CMS and would be released by early 2007, followed by another public comment period. However, the Center for Medicine and Medical Services (CMS) abruptly changed its course within the next month and at the Senate hearing on DTC took the position that genetic testing is already covered under CLIA and that a specialty area would only affect analytic validity and would fail to address the larger concerns about Direct-to-Consumer Marketing (DTC), namely marketing, sales, and interpretation and communication of results.

Direct-to-Consumer Marketing: In late July, the Senate Special Committee on Aging had a hearing based on a Government Accountability Office (GAO) investigation of direct-to-consumer marketing of genetic tests. The investigation reported that four different companies offering "nutrigenomic" testing over the internet provided misleading and scientifically questionable results, and then urged customers to use expensive, "personalized" nutritional supplements that some of the companies also offered for sale. Consumers should be made aware of this activity.

Three panels of witnesses testified at the hearing: the GAO and Kathy Hudson from the Genetic and Public Policy Center; representatives from the DTC testing companies; and representatives from the Center for Medicare and Medicaid Services and the Food and Drug Administration. The FDA took a strong position against the DTC tests and when pressed by the Chairman, stated that the agency believes it has the authority to regulate such tests. FDA plans to oversee and regulate tests done in-house. Warning letters were sent to a few companies, stating FDA's intention to watch them closely and develop new regulations. However, there is concern that a heavy-handed effort could slow down legitimate studies.

Council discussed FDA regulation of the direct-to-consumer genetic tests and newborn screening.

PROJECT UPDATES

POPULATION GENOMICS (GAIN/GEI)

Dr. Teri Manolio, Senior Advisor to the Director for Population Genomics, provided an update on the Genetic Association Information Network (GAIN) and Genes and Environment Initiative (GEI).

GAIN is a public-private partnership led by the Foundation for NIH (FNIH), which involves NIH, several corporations, private foundations, advocacy groups and concerned individuals. Through the whole-genome genotyping of samples from existing case-control studies, the project will contribute key data to the effort to identify genetic contributors to disease risk. A key feature of GAIN is that the data will be made rapidly available for free access by qualified members of the scientific community, while maintaining participant confidentiality. As of September 2006, GAIN is in its final project selection process. Peer review was organized by the FNIH, and was followed by analysis of the proposed projects by the GAIN Technical Advisory Group (TAG). On the basis of the results of peer review and the subsequent technical analysis, recommendations were made to the GAIN Steering Committee as to which studies offer the best chances of finding gene regions contributing to diseases or traits.

The TAG members include staff and investigators from NIH, FNIH, MIT/Broad Institute, University of Michigan, University of Pittsburgh, University of Wisconsin, Perlegen and Pfizer. During the review process, four TAG subgroups have worked on sample ascertainment, genotyping, consent/IRB, and power and analysis, with each group concentrating on its specific issue in each application. The sample ascertainment subgroup focused on epidemiologic design, biases in selection of cases and controls, validity of phenotypic characterization and extensiveness of shared data. The genotyping subgroup, including investigators from Perlegen and Broad Institute, assessed the quantity and quality of DNA needed. The consent/IRB subgroup focused on restrictions on data use and adequacy of the consent process. The power and analysis subgroup discussed sample size (suggested modifications), analysis plan (associations, population substructure) and plans for replication.

The TAG process uncovered several challenges, the biggest being restrictions on data use in the existing consents. GAIN policy is that data will be made available to the scientific community for free and open access, but some of the studies that did well in peer review included consent forms that specified that data would not be used for commercial use. For this reason, such studies could not be used. Fortunately, a large enough number of studies without these restrictions scored well enough to meet GAIN's needs. There were also problems from inadequacies, inconsistencies, or unrealistic commitments in the consent process, such as a statement that upon notification from a submitter, researchers would remove a sample from a study within 48 hours. Other technical challenges included major inconsistencies and incompleteness of submitted datasets (in spite of what had been described in the application), and previous whole genome amplification or other inadequacies in submitted samples. Those studies with the potential for doubling sample size or extending to population samples of different ancestry were very appealing.

The FNIH Board will make its decisions by mid-September and genotyping will begin in October. The TAG will work with the investigators of the chosen studies to optimize the design of each study (e.g., which samples will be chosen for genotyping, which platform) and help to identify other potential improvements. An analysis workshop for the study Principal Investigators and primary analysts will be held on November 29-30 in Bethesda. Genotype-phenotype data should be available in January or February 2007, with a general 4-month turnaround for genotyping.

Dr. Collins explained that there was initially an expectation that 14,000 samples could be analyzed by the amount of funding available to GAIN, but he noted that with technology development and cost reduction, that number may increase to around 18,000. There are enough highly scored studies to gather 18,000 samples.

Council members offered comments on the problem of withdrawing samples upon request from the donors and the advantages of the TAG process for detecting major challenges in applications. Council also asked whether the chosen applications would require reconsent of participants, since many consent forms were not developed with data release in mind. Dr. Manolio responded that most studies had adequate consent procedures and met the standards to place samples into a controlled access database. She also noted that one of biggest challenges is that different studies do not report phenotypes, even some that would be expected to be straightforward, such age or sex, in the same way. She said that there it would be opportune to develop and use standard ontologies across studies, and suggested that NIH could start this effort, by choosing a small number, say 10, key variables and create standard definitions for them.

Council suggested that once the GAIN project is underway, if a particular issue were not being studied, NIH could provide supplements or fund additional applications to study something very specific. All agreed that many more research ideas will be generated once the GAIN data are released.

Dr. Manolio then turned to the Genes and Environment Initiative (GEI). Funding for GEI is included in the President's FY07 budget but, as noted earlier, the 2007 appropriation has not

been passed yet. However, NIH is going ahead and releasing RFAs to solicit proposal for GEI now so as to be ready to spend the GEI funds in 2007 should they be made available. GEI aims to accelerate understanding of genetic and environmental contributions to health and disease. There are two components to GEI.

The first is the genotyping of case-control studies of common disease (proposed at $26M per year for four years). The second is development of innovative technologies to measure environmental exposures, diet, and physical activity (proposed at $14M per year for four years). An NIH-wide Coordinating Committee, with subgroups responsible for each of the components, is leading the project.

Council asked about the plans for following up the genome-association studies. Dr. Manolio answered that GEI is being planned to include follow-up replication, fine-mapping, sequencing, functional, and translational studies. She provided a breakdown of the proposed FY07-FY10 budget that included specific amounts for: GWA studies, data analysis, replication/fine-mapping, sequencing, database, functional studies and translational studies. The total FY07 expenditures will actually be greater than $26M, as the sequencing component is being donated by NHGRI. GWA and data analysis require the largest percentage of the budget. Currently, replication studies are expected to require 7% of the budget, but this may have to increase later. Responses to the RFAs may include replication and fine-mapping, either as follow-ons to GWAS analysis or to follow up on GWA studies that have already been completed. $8.7M remains uncommitted to be used for additional GWA, data analysis, fine mapping, functional studies or translation research as needs arise.

The GEI Requests for Applications, HG-06-014 (genotyping facilities), HG-06-032 (Data Coordinating Center) and HG-06-033 (study investigators; sample sharing and analysis), were released on September 7, 2006. The RFAs can be found at http://grants.nih.gov/grants/guide/rfa-files/RFA-HG-06-033.html. GEI expects to support analysis of about 15 complex diseases or traits altogether, with selection being made in two or three rounds. Investigators may apply for initial discovery genotyping, replication genotyping, or both. The Data Coordinating Center will provide analytic support, data quality assessment and quality control, and logistical management of the GWA program. Funds will also be provided to investigative groups to support the submission of samples from well-characterized subjects for GWA genotyping and/or replication studies, and to analyze the resulting data. The program will promote standardization and harmonization of phenotypic and environmental exposure data to permit cross-study analyses.

Letters of intent for the genetics components of the GEI are due November 1, 2006 and the application receipt date is November 29, 2006. The applications will be reviewed in March 2007. The GEI Genetics Subcommittee and Coordinating Committee Review will take place April 26, 2007, and the applications will be reviewed by the NACHGR at its May 2007 meeting. The start date for the GWA component is July 2007. The project will run for four years. Eight or nine disease studies will be supported in the first round of the

project. An RFA will be reissued in FY08 to solicit additional studies for a second round, and a third round will be held in FY09. Annual workshops on analysis of GWA data are planned for October 2007-2009. These workshops will be open to a variety of investigators. Additional RFAs, for analysis of sequence data and analysis of gene-by-environment interactions, are anticipated. One workshop, on design of sequencing studies to follow up GWAS, is planned for March 2007. The first round of sequencing projects will follow the first GWAS, and will be selected by both the GEI process and the NHGRI medical sequencing process.

Functional and translational studies are scheduled for later in the GEI timeline, as they will based upon data generated earlier in the Initiative. The plan is to move to functional studies after GWAS and sequencing. Another workshop is being planned to consider appropriate approaches to functional genomics. As mouse knockouts are one important approach to functional analysis, GEI plans to provide funds to the KOMP to knock out the orthologues of candidate genes identified in the GEI studies. Two additional RFAs are planned to provide support for translational studies of identified variants.

Council urged Staff to engage the community and provide a description of the opportunities available through published articles. Council also asked whether the problems in the applications identified by the GAIN TAG process had changed anything in the planning for GEI. Dr. Manolio answered that submission of datasets will be required before the peer review. The integrity of the data will be part of the criteria evaluated by the Scientific Advisory Board.

ROADMAP 1.5

Dr. Collins discussed a new process that is now underway to identify potential projects for a second round of funding through the NIH Roadmap. The funds available for new projects are limited, and the next round of funding will come from turnover of Roadmap resources. The Roadmap budget is now planned to be $500 million by FY2008, and a 10% turnover each year is anticipated giving $50M a year to begin new initiatives. For the next round (termed Roadmap 1.5), a set of criteria have been defined for a Roadmap project. For a project to qualify as a Roadmap initiative, it must be truly transforming for biomedical research, it must engage participation of NIH as a whole and its outcomes must synergistically promote the missions of many or all of the Institutes and Centers, and it must be a project that no single entity is likely to do.

Three sources of input have been established for Roadmap 1.5. The first will solicit ideas from meetings of outside consultants, the second will request suggestions from the Institutes and Centers, and the third will solicit input from the broad stakeholder community. For Phase 1, five meetings were to be held, each with about twenty external consultants representing a broad range of expertise. Each of the meetings is co-chaired by two Institute/Center directors. Three of the meetings have already occurred; the September 21st meeting will be co-chaired by Dr. Collins and Dr. David Schwartz (NIEHS). The participants were asked to submit, prior to the meeting, three projects for discussion, and the objective is to come to agreement on three to five projects to forward for further consideration. Dr. Zerhouni attends all meetings. In Phase

2, each of the Institutes and centers were asked to submit one-page descriptions of up to 5 ideas. These proposals were due in August and will be added to the suggestions from the consultant meetings. NHGRI submitted five proposals. Phase 3 will consist of a Request for Information asking for input from the community on the projects compiled in the first two phases, as well as additional ideas; individuals may submit up to three of their own ideas. All of the proposals and comments received will be reviewed for their responsiveness to the Roadmap criteria by NIH staff. Institute and Center directors will then meet to discuss the initiatives and to choose a number to pursue. Ultimately, there will be Requests for Applications or Requests for Proposals for funding in FY2008.

Dr. Collins encouraged Council members to look at the Request for Information in October, comment on the proposals, indicate their preferences, and suggest additional ideas. It is expected that many proposals will be highly relevant to NHGRI's agenda and the Institute is looking forward to the outcome of the process.

Council asked if there is a relationship between these initiatives and what Congress is calling a Common Fund in the NIH Reauthorization. Dr. Collins noted that Congress is working on reauthorizing NIH for the first time in 13 years. One aspect of reauthorization is to identify a Common Fund, which is essentially a continuation of the Roadmap process. The NIH Director would then develop processes to decide how to use the Common Fund. Congressman Joseph Barton (R, Tx), chair of the House committee considering NIH reauthorization, thinks the fund should ultimately be 5-10% of the NIH budget. The Common Fund is proposed to start at the current level of the Roadmap and increase to its final size in parallel with future growth of the NIH budget.

NHGRI SCIENTIFIC PRIORITIES

Dr. Guyer presented a status report on NHGRI scientific priorities. In a brief background, he described the development of NHGRI's "Vision for the Future of Genomic Research" in 2003. He pointed out that that document was a statement of opportunities but did not address the issue of setting priorities among the opportunities enough to comprise a plan. Since 2003, NHGRI has faced the challenge of implementing the ideas in that statement and has found the issue of priority-setting to be increasingly difficult. During the past summer, Staff revisited the 2003 "vision" document, and asked several questions. Are the 2003 conclusions still relevant? What has already been accomplished or completed? What new opportunities need to be taken into account by NHGRI and how should they be implemented? How much will these cost?

Staff members reviewed each of the Grand Challenges devised in 2003, addressed the activities that remain to be finished, outlined new ones, and estimated the cost and priority of each. Staff concluded that the Grand Challenges were still a useful way to look at NHGRI's goals for the next several years. A tremendous amount of progress has been made since 2003 toward addressing them; examples include the sequencing of 24 mammalian genomes to coverage of 2X or greater and the sequencing of many other genomes, including clusters around important model organisms; continued improvement in sequencing technology; the isolation of large numbers of human, mouse and rat full-

open reading frame cDNAs by the Mammalian Gene Collection (MGC) project; the HapMap and its application to the analysis of the genetics of complex disease; the ENCODE pilot project; the NIH Roadmap's Molecular Libraries Initiative; new policies in the areas of intellectual property and data release to encourage widespread use of all of genomic data; and many others. At the same time, almost none of the Grand Challenges have been completed and there is still an enormous amount of work remaining.

Identifying priorities among remaining activities proved challenging. The estimated cost for funding all of the activities considered to be of high priority substantially exceeds NHGRI's expected budget. While priority setting is difficult, it is also essential. Dr. Guyer described the goal of create a working set of present priorities that can subsequently function as a framework to evaluate new ideas that are developed. As new ideas arise, they can be compared against the existing set of priorities to inform decisions on how to move forward to most effectively address NHGRI's mission of improving human health through genomics.

Dr. Guyer presented an analysis of NHGRI's spending on the Grand Challenges through FY2005, the last year for which the complete data are available as of September 2006. The bulk of NHGRI's funds has been spent in the area of Genomes to Biology, and the majority of those funds has been spent on large-scale sequencing. NHGRI continues to make unique contributions in this area. However, the fraction of the annual budget spent on production sequencing has been decreased, starting in FY05, to make funds available for other programs. Council was reminded that the NHGRI budget has been either flat or reducing for the last few years with increased inflation.

Dr. Guyer then discussed the on-going activities of the NHGRI's extramural program and new areas of interest, including continued large-scale sequencing for purposes of comparative genomics and medical sequencing, including The Cancer Genome Atlas (TCGA) project; technology development to reduce sequencing costs two orders of magnitude (the "$100,000 genome") and then another two orders of magnitude (the "$1,000 genome); activities in the broad area of genomic function, including the scaling up of ENCODE, modENCODE, completion of MGC, the Knockout Mouse Project (KOMP), and proteomics; translational research, including population genomics and chemical genomics (opportunities in both of these areas were recognized in 2003, but only recently have NHGRI research programs started to materialize); and the ELSI, CEGS, Minority Action Plan, training, and SBIR/STTR programs.

In their analysis, Staff attempted to estimate how much funding would be required for NHGRI to make significant contributions in each of these areas. Overall, support for all of the activities that Staff identified as very high or high priority in FY07 would require $60 million more than the than the proposed budget for FY07. Dr. Guyer then asked Council to provide guidance on setting priorities.

Council noted the difficulty of the decisions being faced by NHGRI. In the ensuring discussion, attention was directed to the ENCODE Project, which was recognized as a high priority. The question was raised whether ENCODE is ready for a substantial ramp

up this year, and Council asked if there is a process for making that decision. Dr. Guyer responded that the scale up of ENCODE was initially scheduled for 2006, but that an assessment of the progress of ENCODE, including input from a workshop with the ENCODE Scientific Advisory Panel and additional expert, concluded that the program was not quite ready to scale up. Subsequently, the Council approved the extension of the pilot projects for an additional year so that both Staff and current participants would have additional time to address issues such as definition of data standards, production metrics, and cost analysis. Council recommended that the RFA for scale up be written very clearly and explicitly with respect to objectives and expectations for the scale-up phase, and to define stringent review criteria and then to hold the proposals to a high standard. Plans for scaling up ENCODE to analyze the entire genome call for an increase in annual funding from $10 million to $25 million in the first year, with $5 million increases annually thereafter. Council made it clear that, in this time of budgetary stress, if no applications meet the standards for scale up, NHGRI should not fund any scale-up. However, as Council still considers ENCODE to be a very important initiative, the pilot efforts may be have to be continued.

Council then turned its attention to modENCODE and expressed some concern about the relative timing of the new initiative with respect to ENCODE. If the scale up of ENCODE is being delayed because important questions about whole genome analysis of sequence-based functional elements cannot be addressed yet, is it premature to go ahead with modENCODE? Dr. Guyer answered that, while the issues that confront the analysis of a 100Mb genome may or may not be the same as those that face the analysis of a 3000 Mb genome, Staff is confident that responses to each initiative can be evaluated on their own merits, and that the decision to delay the scale up of ENCODE should not affect the evaluation of modENCODE proposals. However, modENCODE proposals will be evaluated with the same stringent standards being developed for the ENCODE scale up.

Council suggested that the timing may work such that modENCODE may inform ENCODE. It is not simply the size of the genome, but the number of elements (e.g., transcription factor binding sites) and biological conditions (e.g., tissues or developmental stages) will differ in these projects. Much of the analytical capacity will be determined by reagent availability and quality control issues. Lessons learned from modENCODE may have a positive effect on how to execute the scale up for the human genome. Another issue raised by Council was cost analysis, noting that currently in ENCODE, there is no clear approach to cost analysis by the data producers and little reliable estimate on the magnitude of potential cost decreases. Dr. Guyer agreed, saying that Staff is worried about the cost issue, recognizing how important a clearly defined approach to cost analysis was to the sequencing program and noting that ENCODE has not been able to develop a comparable analytic approach yet. This is another issue that applicants to both the modENCODE RFA and later the ENCODE scale up RFA will be asked to address in detail.

Council asked if there would be enough time between the funding of modENCODE and the human ENCODE to provide information to help the ENCODE scale up as it appears

that the timing may be too tight for this. Dr. Guyer responded that the timing is tighter than ideal, but there will be ongoing development in both projects.

Council suggested that NHGRI may need to re-think the ENCODE Project, and consider what aspects of the project can achieve the economies of scale and which might not. It was noted that one of the conclusions of the assessment meeting in January was that large-scale projects are only valuable when they can achieve lower cost and higher quality for the same work than can be achieved by single laboratories. Council suggested that NHGRI look into the possibility that large amounts of transcription factors antibodies could be produced in a large-scale, cost-effective manner. Council further suggested that NHGRI assess which aspects of ENCODE belong in a large program and which should be done as individual projects. It does make sense to identify elements that can be analyzed on a "dollar in, element out" and that those should be catalogued at a large scale. For other elements for which this cannot be done, switching to R01 support could have the additional benefit to NHGRI of distributing the costs across NIH.

On another topic, Council asked what funding the Genes and Environment Initiative (GEI) gets from the NHGRI budget and whether medical sequencing contributes to it. Dr. Collins responded that the NHGRI contribution of sequencing capacity to GEI will be considered to be part of the budget for medical sequencing. Initial calculations estimated that the amount of sequencing needed for GEI was relatively small and could be completed for $3.5 million total over several years. As a result NHGRI decided to absorb the costs rather than take funds from the GEI pot. Dr. Guyer added that these are the kinds of projects NHGRI expected to fund in the medical sequencing program anyway.

Dr. Collins stated that if NHGRI had continued a steady 5-6% funding growth rate since 2003, NHGRI's budget would be right where it needed to be. However, as this had not happened, NHGRI is facing a circumstance where it has to make very difficult decisions, and that there will be projects that would make important contributions to medical science but cannot be pursued within NHGRI's budget. Council concurred that NHGRI has to continue to get the message out that it is limited by resources, not by ideas or talents.

Council asked how much is being spent on the technology development program to decrease sequencing costs to the $1K genome. Dr. Guyer answered $25 million per year.

COUNCIL-INITIATED DISCUSSION

During February Council, there will be a presentation from the intramural program. The modENCODE proposals, applications for ENCODE technology development and CEGS applications will be reviewed.

Council asked for an update on the CEERs Program and the larger ELSI program. The CEERs is currently starting its third year of funding, and an RFA will be released for the second round of centers. Site visits will take place this spring, and staff suggested that

May Council would be better timing for a CEERS update as well as a presentation on the ELSI program. Council concurred.

Staff noted that NHGRI manages a Roadmap center in the Bioinformatics and Computational Biology Initiative on the development of tools to develop and store ontologies. This has recently been an area of activity in data integration, and Staff suggested that Council might be interested in a presentation on progress toward the "industrialization" of the production of ontologies. Council concurred

Council suggested that it would be interested in a presentation about the NHGRI R01 portfolio, and that getting feedback from individual R01 investigators may be interesting.

Council also expressed interest in a presentation on new sequencing technology and the reality of the $100K and $1K genomes.

ANNOUNCEMENTS AND ITEMS OF INTEREST

Dr. Guyer directed Council to the Council folders containing a sampling of the kinds of discussions that have been in popular press.

CONFLICT OF INTEREST

Dr. Guyer read the Conflict of Interest policy to Council and asked them to sign the forms provided.

REVIEW OF APPLICATIONS

In closed session, the Council reviewed 122 applications, requesting $206,245,714. The applications included 38 regular research grants, 7 pilot projects, 2 program projects, 11 ELSI grants, 34 RFA grants, 1 area grant, 3 center grants, 1 conference grant, 1 continuing education training program grant, 13 SBIR Phase I grants, 4 SBIR Phase II grants, 4 fellowship grants and 3 others. A total of 73 applications totaling $194,392,541 were recommended.

I hereby certify that, to the best of my knowledge, the foregoing minutes are accurate and complete.

2/20/07
Date

Mark Guyer, Ph.D.
Executive Secretary
National Advisory Council for Human Genome Research

2/20/07
Date

Francis S. Collins, M.D., Ph.D.
Chairman
National Advisory Council for Human Genome Research