NATIONAL ADVISORY COUNCIL FOR HUMAN GENOME RESEARCH
SUMMARY OF MEETING[1]
May 21, 2007

The open session of the fiftieth meeting of National Advisory Council for Human Genome Research was convened at 8:36 A.M. on May 21, 2007 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 8:36 A.M. until 3:05 P.M. on May 21, 2007. In accordance with the provisions of Public Law 92-463, the meeting was closed to the public from 3:15 P.M. on May 21, 2007 until adjournment for the review, discussion, and evaluation of grant applications.

Council members present:

Eric Boerwinkle
Andrew Clark
Jorge Contreras
Marilyn Coors
Geoffrey Duyk
Sean Eddy
Vanessa Northington Gamble
Richard Gibbs
Caryn Lerman
Deidre Meldrum
Patrice Milos
Jeffrey Murray
David Page
Stephen Prescott
Harold Shapiro
Lincoln Stein
Paul Sternberg
Richard Weinshilboum

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".

Ajay, DER
Catherine Bennet, DER
Saveri Bhattacharya, DER
Christianne Bird, DER
Vivien Bonazzi, DER
Vence Bonham, OD
Joy Boyer, DER
Lisa Brooks, DER
Comfort Browne, DER
Ernsly Charles, DER
Debbie Chen, DER
Cheryl Chick, DER
Monika Christman, DER
Francis Collins, OD
Karen DeLeon, OD
Carla Easter, OD
Janelle Everett, OD
Elise Feingold, DER
Adam Felsenfeld, DER
Colin Fletcher, DER
Phyllis Frosst, OD
Barbara Fuller, OD
Peter Good, DER
Bettie Graham, DER
Eric Green, DIR
Alan Guttmacher, OD
Mark Guyer, DER
Linda Hall, DER
Sarah Harding, OD
Emily Harris, OD
LaTasha Harris, OD
Chris Juenger, DER
Bill Kibby, OD
Laura Liefer, DER
Carson Loomis, DER
Teri Manolio, OD
Jean McEwen, DER
Keith McKenney, DER
Jessica Melone, DER
Ken Nakamura, DER
Vivian Ota Wang, DER
Kenneth Ow, OD
Brad Ozenberger, DER
Carmen Perera, DEAS
Jane Peterson, DER
Rudy Pozzatti, DER
Eddie Rivera, OD
Jerry Roberts, DER
Cristen Robinson, DER
Anna Rossoshek, DER
Jeff Schloss, DER
Geoff Spencer, OD
Jeff Struewing, OD
Carolyn Taylor, DEAS
Gary Temple, DER
Elizabeth Thomson, DER
Larry Thompson, OD
Susan Vasquez, OD
Lu Wang, DER
Kris Wetterstrand, DER
Diane Williams-Bey, DEAS

Others present for all or a portion of the meeting:

Diane Baker, Genetic Alliance
Joann Boughman, American Society of Human Genetics
Camilla Day, CSR
Judith Benkendorf, American College of Medical Genetics
Norma Kim, RTI
Sharon Olsen, International Society of Nurses in Genetics
Branka Sekis, Social and Scientific Systems, Inc
Sharon Terry, Genetic Alliance
Wendy Uhlmann, National Society of Genetic Counselors

INTRODUCTION OF NEW MEMBERS AND STAFF, LIAISONS AND GUESTS

Dr. Guyer introduced nine new Council members: Eric Boerwinkle from the University of Texas Health Center at Houston, Jorge Contreras from WilmerHale, Richard Gibbs from the Baylor College of Medicine, Caryn Lerman from the University of Pennsylvania, Patrice Milos from Pfizer, David Page from the Whitehead Institute for Biomedical Research and Massachusetts Institute of Technology, Lincoln Stein from Cold Spring Harbor Laboratory, Paul Sternberg from the Howard Hughes Medical Institute California Institute of Technology, and Richard Weinshilboum from the Mayo Clinic College of Medicine. Dr. Guyer also noted the departure of two Council members: Mary Hendrix and George Weinstock.

Dr. Guyer introduced new NHGRI staff: Mary Affeldt, Executive Officer; Janelle Everett, Management Analyst; Tasha Harris, Program Assistant in PPAB ; Bill Kibby, Deputy Chief Information Officer; and Jeff Struewing, Program Director for Population Genomics.

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

APPROVAL OF MINUTES

The minutes from the February 2007 Council meeting were approved as submitted.

FUTURE MEETING DATES

The following dates were proposed for future meetings: September 10-11, 2007, February 11-12, 2008, May 19-20, 2008, September 8-9, 2008, February 9-10, 2009, and May 18-19, 2009.

DIRECTOR'S REPORT

I. GENERAL ANNOUNCEMENTS

Mary Affeldt joined NHGRI in April as the new Executive Officer. Prior to her arrival at NHGRI, she was the Branch Chief, Administrative Management Branch of the Intramural Research Program at the National Institute of Drug Abuse (NIDA).

On April 2, 2007, Barbara Alving, M.D., was named as the Director of the National Center for Research Resources (NCRR) subsequent to her position as Acting Director of NCRR where she oversaw the launch of the Clinical and Translational Science Awards (CTSA) program.

Griffin Rodgers, M.D., was appointed as the Director of the National Institute on Diabetes, Digestive, and Kidney Diseases (NIDDK) on April 1, 2007. Dr. Rodgers is widely recognized for his work on the genetics and therapy of hemoglobin disorders and for his contributions to the first effective and FDA-approved therapy for sickle cell anemia.

Stephen Straus, M.D., the first Director of the National Center for Complementary and Alternative Medicine (NCCAM) passed away on May 24, 2007, after a long battle with brain cancer. Dr. Straus was a highly recognized researcher in the area of herpes viruses and immunological diseases.

Eric Green, M.D., Ph.D., was elected to the Association of American Physicians. Dr. Green was one of 44 new members elected to AAP in 2007. At NHGRI, Dr. Green holds several positions including Scientific Director, Chief of the Genome Technology branch, and Director of the NIH Intramural Sequencing Center (NISC).

Council member Jeff Murray, M.D., was honored with the Curt Stern Award for Human Genetics.

The National Academy of Sciences elections in 2007 include a number of individuals with research interests closely aligned to NHGRI: David Ginsburg, Investigator, Howard Hughes Medical Institute, and James V. Neel Distinguished University Professor, University of Michigan Medical School, Ann Arbor; Helen H. Hobbs, Investigator, Howard Hughes Medical Institute, and Director, McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas; Brian J. Druker, Investigator, Howard Hughes Medical Institute, and JELD-WEN Chair of Leukemia Research, Oregon Health & Science University, Portland; Sean B. Carroll, Investigator, Howard Hughes Medical Institute, and Professor of Molecular Biology and Genetics, University of Wisconsin, Madison; and Jonathan G. Seidman, Henrietta B. and Frederick H. Bugher Professor of Cardiovascular Genetics, Harvard Medical School, Boston.

The 2007 Public Health Leadership Award from the National Organization of Rare Disorders was awarded to William Gahl, M.D., Ph.D., who serves as the Clinical Director for the National Human Genome Research Institute.

NIH Director Elias Zerhouni, M.D., launched the New Innovator Award, funded through the NIH Roadmap Common Fund, which is available to investigators who have not yet obtained an R01 award but who could have a great impact on biomedical or behavior science. The application period opened on April 25 and will close on May 22, 2007. NIH expects to make at least 14 awards in September 2007 in grants of up to $1.5 million in direct costs over five years.

II. NEW NHGRI INITIATIVES

The Office of Population Genomics has released two RFAs: "Genome-Wide Studies in Biorepositories with Electronic Medical Record Data" and "High-Priority Phenotype and Exposure Measures for Cross-Study Analysis in Genome-Wide Studies." A web-based pre-application information meeting was held for the two RFAs, applications for which were due May 17, 2007.

NIH is accepting pre-applications (X02) for the probe development center component of the next phase of the Molecular Libraries Roadmap initiative. The goal of the Molecular Libraries Initiative is to provide public-sector small molecule screening capabilities for the purpose of developing new research probes. After a very interesting and successful pilot phase, the probe development program is ready to scale up. The external scientific advisors are enthusiastic about the program, and after several presentations to the NIH IC directors, the program is moving ahead with new RFAs to solicit the participants in the production phase. An X02 pre-application must be submitted in order to be invited to submit a full application. The pre-applications for both the Molecular Libraries Probe Production Center Network and the Cheminformatics Research Centers initiative are due June 28, 2007 [N.B. The Cheminformatics Research Centers X02 initiative was subsequently withdrawn].

III. RECENT SCIENTIFIC ACCOMPLISHMENTS AND ISSUES

EXTRAMURAL PROGRAM

A summary of the traces in GenBank for all organisms for which NHGRI is supporting genome sequencing can be found in the NHGRI Sequencing Update table. New species were not "approved" for sequencing in the most recent round of sequencing project selection; instead, the working groups prioritized previously approved organisms.

The rhesus macaque genome was published by the Rhesus Macaque Genome Sequence and Analysis Consortium, which is funded in part by NHGRI. The sequencing was done jointly by the Baylor College of Medicine Human Genome Sequencing Center in Houston, the Genome Sequencing Center at Washington University of Medicine in St. Louis, and the J. Craig Venter Institute in Rockville, MD. The researchers have identified 200 genes likely to be key players in recent primate evolution.

Researchers at the Broad Institute of MIT and Harvard and its collaborators published the first genome of a marsupial, *Monodelphis domestica*, in the journal *Nature*. The analysis of the marsupial genome gives researchers information on a critical part of the evolutionary tree. The comparison of the marsupial genome to the genomes of non-marsupials reveals that most innovations leading to the human genome sequence do not lie in protein-coding sequences.

This year's Cold Spring Harbor Laboratory "The Biology of Genomes" meeting was a very intense 3.5 days, during which many interesting and important developments were reported. Two prominent themes were clear: the identification of genetic variations associated with common disease and the emergence of the new sequencing platforms (454/Roche and Illumina/Solexa). It is now possible for one sequencing run to generate as much as 1.5 GB of raw sequence data, and researchers will now be challenged by the need to deal with tera- and

even peta-bases worth of data. One example of the way in which the new sequencing technologies are already having a significant effect is that ChIP-chip methodology is being supplanted by the use of high throughput short read sequencing to analyze the immunoprecipitated material ("ChIP-seq"). The ELSI session of the meeting, which NHGRI helped to organize, focused on the rapidly accumulating evidence for selection in the human genome. On the last night of the conference, two guest speakers, Eddy Rubin and Tom Hudson, highlighted DOE/JGI sequencing efforts and the projects and plans of the Ontario Institute for Cancer Research, respectively. Additionally, Jim Watson gave an impromptu talk regarding the sequencing of his own genome. On May 31st, a press conference will be held in Houston to announce that Dr. Watson's genome has been sequenced to 6-fold coverage. Among the several satellite meetings held at Cold Spring Harbor was the International Sequencing Consortium. Issues of shared concern were discussed, including which organisms various groups are sequencing, new technologies, and data release.

A Human Genome Reference Consortium has been established to address long-term curation of the human genome reference sequence. The NACHGR had recommended that the Institute organize this effort because even though the human sequence is completed, there are still gaps that are going to be filled by new technologies and corrections that will need to be made. The Human Genome Reference Consortium consists of NCBI, EBI, WUGSC and the Sanger Institute and has appointed Richard Gibbs as chair.

NHGRI has funded a new sample repository at the Coriell Institute. As its interests have developed and its needs have grown, NHGRI decided to establish its own sample repository. The repository will include all of the samples collected for the International HapMap Project, except the CEPH samples from Utah. In addition to the initial HapMap populations (Yoruba, Han Chinese, and Japanese), the NHGRI repository includes samples from seven additional populations: Luhya (Kenya), Maasai (Kenya), Gujarati (India, in Houston), Chinese (Denver), Mexican origin (LA), African ancestry (US Southwest) and Tuscan (Italy). Over the next few months, these samples will be studied to assess the applicability of the HapMap to additional groups. The Coriell Institute will follow same process of community engagement that was used for the initial HapMap populations. No identifying information or connections of genotype and phenotype information are included. Community advisory groups, which serve as a liaison between people in the donor community and the repository, have been established for all of the sets of samples.

In addition to partnering with NCI for The Cancer Genome Atlas (TCGA), NHGRI has been supporting the Tumor Sequencing Project, which serves as a pre-pilot for the TCGA. The consortium has completed the sequencing of ~900 genes in 188 lung adenocarcinoma samples. These samples, plus an additional set of 200, were analyzed for copy number changes. Several new candidate loci were identified.

The TCGA pilot project is underway with characterizations of the first specimens of glioblastoma multiforme. The sample collection aspect of the project has proven to be challenging. The sequencing and genome characterization centers are ramped up, but the rate-limiting step at present is sample acquisition, given the very high standards for tumor purity and consent that the TCGA pilot has adopted. This issue has been particularly difficult for

glioblastoma because of necrosis is one of the features used to characterize this tumor type. Ovarian and lung carcinoma tissue repositories have now been identified, and samples will enter the pipeline later this year. NCI is pushing hard to find new sources to keep up with the genome centers' demand for analytes.

The Encyclopedia of DNA Elements (ENCODE) Project has a paper in press in a high-profile journal that describes the findings of the ENCODE pilot project, focused on 1% of the human genome (30 Mb). In light of the success of the pilot project, ENCODE is currently scaling up to full production phase, the applications for which will come to September Council. After the inception of the ENCODE pilot, the NACHGR recommended that a parallel study, looking at all functional elements in model organisms would be highly beneficial. This has now led to the establishment of the modENCODE (Model Organism ENCODE) project to identify comprehensively functional elements in the genomes of *Caenorhabditis elegans* and *Drosophila melanogaster*. The first meeting of the modENCODE consortium was held in May, just prior to the Biology of Genomes meeting at Cold Spring Harbor.

The Knockout Mouse Project is a trans-NIH initiative that has the goal of systematically knocking out every protein-coding gene in the mouse genome by targeting genes in embryonic stem cells. The library of knockout ES cells will be available to the whole research community. There is considerable international activity in systematic mouse gene knockout efforts, so the NIH, The European Commission, and Genome Canada have formed the International Knockout Mouse Consortium (IKMC), which held its first meeting in Brussels on March 14, 2007. At the meeting, a number of issues were addressed to ensure that the collaboration is effective. The Texas Institute of Genomic Medicine (TIGM) subsequently joined the IKMC, as announced in a letter to the editor of *Cell*. The Knockout Mouse Project and 13 NIH institutes are working together to provide administrative supplements to have mouse knockouts made from existing mutant ES cell resources and deposited into an NIH-supported repository. They have issued one administrative supplement so far. The funds were provided with the understanding that the mouse, once generated, would be cryopreserved and put into the repository.

The Mammalian Gene Collection (MGC) is an effort to generate full-length cDNA clones for all human and mouse genes, with a smaller component for rat and cow. MGC currently has derived and made available 15,680 human genes, 14,827 mouse genes, 5,055 rat genes, and 5,537 cow genes. To reach the MGC goal of one validated clone for each of the ~18,700 RefSeq genes for human and mouse, MGC aims to obtain clones for another 3500-4000 human and mouse genes through DNA synthesis. It is currently projected that the last synthetic clones will be delivered to MGC in mid to late October 2007. One caveat of the project is that the MGC will contain only one representative isoform of all the well-established human and mouse genes. If funds were available, MGC might wish to extend its outlook and collect more isoforms.

The NIH Roadmap for Biomedical Research is engaged in a new competition (Roadmap 1.5) following the initial set of initiatives that were funded in FY2004. In FY08, there will only be about $30M/year available for Roadmap 1.5. Beginning last summer, a systematic process was implemented to identify new Roadmap-appropriate opportunities. On May 18th, the IC

directors met to hear the presentations on the five proposals and three pilot efforts that were the final candidates. One of the proposals that was approved was the Human Microbiome Project (HMP), which aims to systematically analyze the microorganisms of the human body. NHGRI will co-lead this initiative along with NIAID and NIDCR. A proposal for Epigenetics (led by NIEHS and NIDA) and a pilot Connectivity Map proposal (led by NHGRI, NIA, NIAAA, and NIDDK) were also approved.

INTRAMURAL PROGRAM

The annual presentation to Council by Eric Green on DIR took place later in the open session.

Elaine Ostrander, Ph.D., chief of NHGRI's Cancer Genetics Branch, has led an international team of researchers in discovering a genetic variant that is a major contributor to body size in dogs. The findings appear in the April 6, 2007 issue of *Science*.

A US-Finnish team, including researchers at NHGRI and NIDDK, working closely with two other groups, has identified at least four new genetic variants associated with an increased risk of type 2 diabetes and confirmed existence of another six. The findings of all three groups were published in the April 26th online edition of the journal *Science*.

OFFICE OF THE DIRECTOR

In anticipation of the flood of genome-wide association study (GWAS) results that will be emerging shortly, NHGRI held a science writers workshop on May 1 for GWAS. More than 24 leading science reporters were in attendance, including reporters from the New York Times, Washington Post, the Associated Press, Newsweek, *Science*, *Nature*, ABC News, CNN, and National Public Radio. The workshop was held with the hope that the public will receive clear reports that explain the promise and hope of this new approach to research on the genetics of common disorders.

On April 25, the nation celebrated DNA Day 2007 to commemorate the description of the DNA double helix in April 1953 and the completion of the Human Genome Project in April 2003. This year, NHGRI partnered with graduate students in the Biomedical and Biological Sciences Department at the University of North Carolina at Chapel Hill to provide presentations to more than 4,000 high school students in North Carolina. NHGRI also sent more than forty ambassadors to other schools in the Southeast. The traditional DNA Day Chatroom took place on April 25th; over 3000 questions were received, 300 of which were answered. Also, that weekend, the Liz Lerman Dance Exchange presented a workshop titled "Genes in Motion" for high school students from nine Washington D.C. area schools. The workshop combined methods established by NHGRI researchers, science educators, and the Liz Lerman Dance Exchange to teach genetics through movement, spoken word, and creative applications of technology. The NIH Office of Science Education was a partner with Liz Lerman in the development of the workshop

The 2007 NHGRI Short Course, *Current Topics in Genomic Research*, will be held on the NIH campus from August 5 - August 10. This is the Short Course's 10th anniversary and it has developed a diverse curriculum over that time. The 2007 Short Course class is made up of instructors and students from across the U.S. including attendees from several under-represented minority groups.

NHGRI held a meeting for Physician Assistants at NIH in March titled, "Physician Assistant Competencies for Genomic Medicine." Several follow-up activities were identified for NHGRI and the Physician Assistant community to work on together, including holding another meeting in September to monitor progress on integrating genomic medicine into the education agenda for the four national Physician Assistant organizations.

Family history continues to be an area of interest for NHGRI. The Education and Community Involvement Branch will be announcing a Request for Proposals in May for the 2007-2008 Demonstration Project on Family Health History.

NHGRI POLICY

The FY07 appropriation for NIH was passed by Congress as a Joint Resolution for $29,186 million and signed by the President. The FY08 Budget was submitted to Congress on February 5, 2007, before the FY07 budget was finalized, with a request for $28,858 million. The appropriations process for the FY08 budget is underway. The House has already conducted an appropriations hearing, while the Senate decided not to have a single hearing this year, but will instead have five hearings on the NIH so that each institute and center can present information. Dr. Collins testified at the Senate "Frontiers of Science" theme hearings on May 7, 2007.

After being under debate for at least twelve years, the House passed the Genetic Non-Discrimination Act of 2007, H.R. 493, on April 25, 2007, by a vote of 420 to 3. Credit is due to the leadership of the House and Leadership Office, and also to Sharon Terry, the leader of the Coalition for Genetic Fairness. The Senate, which has passed this bill twice before, now must act on the legislation. That bill, S. 358, has already been marked up by the HELP Committee, and is pending a floor vote. It is hoped that the Senate will either pass the House version or negotiate and pass a different version, which will then have to be negotiated with the House. Either possibility is likely at this point. Sharon Terry is very confident that the bill will pass the Senate, although there are still a few issues that could present problems.

In early March, the Laboratory Test Improvement Act of 2007, S. 736, was introduced by Senator Edward Kennedy and was co-sponsored by Gordon Smith in early March. The bill would require manufacturers of laboratory-developed tests to provide the FDA with data on analytical and clinical validity. The bill would also subject all direct-to-consumer tests to regulation by the FDA and would impose labeling requirements.

Senators Obama and Burr introduced The Genomics and Personalized Medicine Act of 2007, S. 976, in March. The bill would establish an Interagency Genomics Working Group to coordinate federal efforts relating to genomic initiatives.

In early February, the Genetic Research and Accessibility Act, H.R. 977, was introduced by Representative Xavier Becerra and Dave Weldon. The one-sentence bill aims to prohibit patents on human genetic material. Congressman Becerra may request technical assistance from NHGRI for input into developing the bill.

The Secretary's Advisory Committee on Genetics, Health and Society (SACGHS) met to discuss the oversight of genetic tests. The meeting also included various presentations relating to gene patents and licensing. They have now released a report on Large Population Studies for review, and are currently requesting public comments on their Draft Pharmacogenomics Report. Both reports can be found at the SACGHS website.

ANNUAL REPORT ON THE DIVISION OF INTRAMURAL RESEARCH

Dr. Eric Green presented an annual report on the Division of Intramural Research (DIR) at NHGRI. The mission of DIR is to develop genetic and genomic approaches for understanding genome function, for identifying the molecular basis of human genetic disease, and for developing effective diagnostic and therapeutic interventions. Forty-five people comprise the research faculty of DIR, including 19 senior investigators, 17 associate investigators, and 9 tenure-track investigators in a broad-based genetics and genomics environment. There are a total of 520 people working in DIR, including 23 Ph.D. students (out of the 400 graduate students contributing to NIH intramural programs). The 520 staff are housed in nine buildings in four cities across Maryland.

In FY07, 18% of the DIR budget is going to the NIH Central Management Fund and Clinical Center tap and 18% is indirect costs, leaving 64% of the budget for operating funds. The research infrastructure of DIR includes seven branches, seven cores, and 3 centers (NIH Intramural Sequencing Center (NISC), NIH Chemical Genomics Center (NCGC), and the Center for Inherited Disease Research (CIDR)). Organizational highlights of the past year include the appointment of four new branch chiefs, recruitment of three new tenure-track investigators, the full establishment of NCGC, renewal of the NISC sequencing program, the establishment of the Office of Translational Research, and the launching of two clinical genomics initiatives.

Nearly all DIR investigators have been recruited from outside NIH. The program is organized according to an academically-oriented structure. It is "notoriously" collaborative within the NIH, thereby enhancing the landscape of genetics/genomics at NIH. The scientific programs of DIR are very broad and take a "base pairs to bedside" approach to research. DIR gets advice from a Board of Scientific Counselors, which provides general oversight of the intramural program and oversees quadrennial reviews of each branch. Council asked how DIR decides the direction to take when an opportunity to recruit arises. Dr. Green explained that it varies, depending on the recruitment. It starts with advice from the Board of Scientific Counselors and the decision is also influenced by the Branch Chief in whose branch the opening occurs, feedback from the quadrennial reviews, and programmatic considerations. Recently, there has been an emphasis on recruiting physician-scientists.

Council also asked how DIR addresses the issue of the geographical dispersal of its personnel. Dr. Green noted that communications technology is very important. Another device is biannual retreats.

Council asked if new sequencing technologies were being utilized in NISC. Dr. Green answered that NISC has not acquired the 454 system since that technology is not well-suited to NISC projects, but a Solexa/Illumina machine has been ordered. Another new technology, the Kalypsys robot at NCGC, which allows a 1536-well plate format for ultra-high throughput screening, is working remarkably well. Its capacity is high enough so that NCGC can do dose-response curves as part of the primary screen.

Council noted that NIH is a unique place and inquired about how DIR takes advantage of that uniqueness. Dr. Green responded that there is a commitment to strong clinical research and that the program is increasingly using the Clinical Center facilities in programs like ClinSeq. Also, clinical genomics initiatives that are intra-disciplinary and high risk can be done in DIR, but would be more challenging in other settings. Similarly, other opportunities, like NCGC, would be harder to mobilize as quickly in an academic setting.

When asked about DIR's computational abilities, Dr. Green referred to the Bioinformatics and Scientific Programming Core, which provides local access and programming support but also is a conduit to broader clusters, to the 15 member Bioinformatics group at NISC, and the bioinformatics group at NCGC.

In the future, Council would be interested in discussing the unique challenges faced by the intramural program. Dr. Green suggested that Bill Gahl or Les Biesecker could give a talk about DIR initiatives in the arena of translational research, so that Council could hear from clinically oriented researchers. Finally, Council asked in what way having graduate students has changed things at NIH. Dr. Green believes that graduate students have changed the culture of the NIH campus to involve a youthful, motivated, and outgoing crowd.

Council requested summaries and/or direct reports of the Board of Scientific Counselors.

NHGRI REPORT ON NEW INVESTIGATORS

In response to a request from Council at the February 2007 meeting, Dr. Mark Guyer presented information about new investigator awards by NIH as a whole and NHGRI taken from the NIH database. The information represented awards over the nine year period FY 1998-2006. The report defined a new investigator as one who has not previously received an R01-equivalent award (a category that includes the R01, R29 and R37 mechanisms). The report from the database excluded some of the mechanisms NHGRI uses disproportionately, like cooperative agreements.

For all investigators NIH received an average of 19,712 R01-equivalent competing applications per year and awarded an average of 6,324 grants, a 32.1% success rate. For new investigators, the number of first time R01-equivalent applications averaged over the nine-year span was 6,434 applications received and 1,518 grants awarded, a 23.6% success rate. For

NHGRI, the nine-year average for all R01 equivalent competing applications was 120 received per year and 38 awarded, a 31.7% success rate. Finally, the number of NHGRI first-time R01 equivalent applications averaged over the nine year span was 54 per year with 14 grants awarded, a 25.9% success rate. Looking at trends, the success rates for all investigators have been declining across NIH, and success rates are also declining for the first time for R01s.

The success rates for new investigators at NHGRI seem to be slightly higher than the all-NIH rate. NHGRI pays special attention to applications from new investigators. Dr. Guyer also remarked that genomics is a new field which is attracting a lot of young researchers. However, he noted that the number of new investigators supported by NHGRI is still small in absolute terms. Council asked what NHGRI would like to see in these trends. Dr. Guyer stated that up until fairly recently, NHGRI was comfortable with the trends since the Institute had been able to fund all of the applications that the staff believed should be funded. However, with the budget strains starting two years ago, and including this year, this has not been the case and difficult decisions have had to be made.

Council was also curious about whether the size of awards has changed over the years. Dr. Guyer admitted that that had not been not investigated, in part because the number of grants NHGRI awards is relatively small, and grant size fluctuates significantly from year to year, which makes it difficult to know what average grant size means. However, it is true that NHGRI grants tend to be larger than the NIH average.

Council acknowledged that the time has come when NHGRI must consider paying particular attention to translating new discoveries in genomics from labs into the clinic. Since each IC has its own approach, one Council member asked about training of physician-scientists at NHGRI. Dr. Graham stated that the Institute has a pilot for the use of the K23 mechanism (Mentored Patient-Oriented Research Career Development Award). Council suggested that NHGRI consider including the K23 in its K mechanism program

SCIENTIFIC PRESENTATION: New Findings from Genome-wide Association Studies of Common Disease

Dr. Collins spoke on new findings from Genome-Wide Association Studies (GWAS) of common disease. NHGRI has long been interested in developing the capability for using whole genome association approaches in the identification of disease-related genetic variants. This is evidenced by a 1997 publication by Guyer, Chakravarti, and Collins, as well as the 2003 NHGRI Vision for the Future of Genomic Research. In order to conduct genome association studies, a comprehensive, dense catalog of human variants is needed.

Genome-wide association studies are predicated on the correctness of the common disease-common variant hypothesis. If genetic variations associated with common disease are instead a collection of rare alleles, one would expect GWAS using common variants to fail in most cases. Until now, however, this has been a fairly untested assertion.

In 2002, a "whole genome association" approach to common disease would have involved the analysis of 10 million common single nucleotide polymorphisms (SNPs) in 1000 cases and 1000 controls. If all DNAs were genotyped for all SNPs, a total of 20 billion genotypes, at 50 cents per genotype it would have cost $10 billion for each disease.

All this has been profoundly altered by the HapMap, the generation of which was led by NHGRI with contributions from many other NIH institutes and other groups around the world, The HapMap is a catalog of common human SNPs and haplotypes that currently includes millions of SNPs along with information on the underlying organization of SNPs into haplotype blocks. By allowing the identification and use of tag SNPs, this profoundly affected the feasibility of doing GWAS by reducing the amount genotyping necessary to ~300,000 tag SNPs for European and Asian populations, and somewhat more than that for African populations.

During the same period, there has been a vast improvement in the performance of the platforms used for high throughput genotyping, which have significantly increased the number of SNPs that can be processed and decreased the cost per genotype. The current cost of a genotype is a fraction of a penny, compared with more than fifty cents before the HapMap Project. As a result, in 2007, a "whole genome association" approach to common disease typically involves the analysis of 300,000 tag SNPs in 1000 cases and 1000 controls, totaling 600 million genotypes at $0.0012 per genotype, or a total of $800,000 per study.

The first HapMap success story was the identification by Josephine Hoh and her colleagues of a single locus (*CFH*) that contributed to age-related macular degeneration. That finding has since been confirmed by many groups. Using the same strategy, another major risk locus has been identified. Together, these two account for ~50% of the risk for disease, and point to powerful new approaches to prevention and treatment. Other GWAS success stories include the identification of *IL23R* as an inflammatory bowel disease gene and the finding of a family of risk variants for prostate cancer on chromosome 8q24.

Dr. Collins then described his own group's work, in collaboration with two other groups, on the project known as FUSION (Finland-United States Investigation of NIDDM), in applying the GWAS approach to type 2 diabetes. From this work, three studies resulted in three papers that were published in April 2007 describing a total of 10 variants that confer increased risks of diabetes. It is not possible to determine accurately the effect that these variants have on the overall risk for diabetes without performing a prospective study with samples collected in an unbiased way. Follow-ups to GWAS-based associations should ideally occur quickly after discovery, which is the basis for the Concept Clearance for Population Genomics that the Council considered later in the meeting.

Given the new window of opportunity to identify hereditary factors in common disease, NIH has formed a public-private partnership with The Foundation for NIH and the private sector (Pfizer, Perlegen, Affymetrix, and Abbott), known as the Genetic Association Information Network (GAIN). The goal of GAIN is to speed up the ability to conduct genome-wide association studies. New Council member Patrice Milos is the leader for the private sector component and Pfizer, Perlegen, Affymetrix, and the Broad Institute have committed to

support the costs of genotyping for six GWAS studies. After peer review and technical evaluation of assessment of sample sets, six studies were chosen last October. Samples from these studies have now entered the genotyping pipeline and the data will begin to appear in June 2007. The data, both genotypes and phenotypes, are submitted to a new database called dbGaP. Though there is no personal identifying information included, access to the data is only available to legitimate investigators, whose applications for access are being reviewed by a Data Access Committee (DAC) chaired by Emily Harris. There will also be a second group, the Data Use Review Board (DURB) that will oversee whether these particular processes are effective in ensuring that data misuses do not occur.

Another HapMap-empowered project is The Genes, Environment and Health Initiative (GEI). This is a Secretarial Initiative for FY07-FY10, and $40M per year has been made available for two purposes: improving technology for exposure biology and a genetics component that involves GWAS, replication, sequencing, functional and translational studies.

Because of the increasing number of GWAS proposals being submitted to NIH, there has been a serious trans-NIH effort to develop a set of policies that can be adopted across NIH for conducting genome wide association studies. The effort is being led by Betsy Nabel, Director of NHLBI. The draft NIH policy for GWAS, which addresses protection of research participants, data sharing procedures, data access principles, and intellectual property use, has been widely circulated for open discussion in the community, and the policy is expected to be finalized sometime this summer.

Dr. Collins concluded by stating that the successes of the HapMap and low-cost genotyping technologies have ushered in a new era of discovery for heritable factors for common disease. Most genes discovered so far would not have been on anyone's candidate list, and many more discoveries can be anticipated in the next 2-3 years. So far, most risk variants are regulatory, not coding, and while individual odds ratios are modest, each new discovery points to a new pathway involved in pathogenesis.

Council noted that a lot of this work is not going to be NIH funded, and questioned whether there were any way to involve the journals more in encouraging authors to submit their information to dbGaP so the data sets are as complete as possible. Dr. Collins noted that a workshop was held to discuss standards that should be applied to studies, and that some journal editors participated in that workshop. Some standards were statistical and some addressed the openness of information, which led to the recommendation that journal editors should not allow publication unless the paper at the very least includes summaries of the data. The ability of investigators to submit individual phenotypes to dbGaP is going to depend on the nature of the consents under which the samples were collected. In any case, Dr. Collins asserted, investigators should submit all allele frequencies for all SNPs in their cases and controls.

PROJECT UPDATE: Population genomics

Dr. Teri Manolio began by reminding Council that the NHGRI has established an Office of Population Genomics to facilitate the application of genomic knowledge to health and promote multi-disciplinary research in epidemiology and genomics. The population genetics program

will support efforts to apply genomic technologies to existing population and clinical studies, and to develop new population resources for investigation of genetic and environmental contributions to complex disease. The program has set five goals:

- establish research resources to identify genes related to complex diseases and their environmental modifiers;
- improve analysis strategies for relating genotypic to phenotypic data;
- build successful NIH-wide collaborations in population-based genomic research;
- develop novel population research approaches; and
- support cross-disciplinary training for geneticists, epidemiologists, clinical researchers, and clinicians.

Two trans-NIH initiatives, GAIN (a public-private partnership) and GEI are the first major activities in this area to get going. GAIN started in February 2006 with a solicitation for well-characterized sample sets to undergo genome-wide genotyping, and the initial choices were announced in October 2006. After a time for sample characterization and other start-up activities, genotyping of HapMap samples was performed on the two platforms selected for GAIN genotyping and then genotyping of GAIN samples began in February 2007. The first genotype-phenotype dataset is expected to be released in June 2007 with subsequent studies completed and released in the six months thereafter. Other activities that address this goal include two RFAs to support the establishment of infrastructure ("Public Consultation to Inform the Design of Large-Scale Studies" and "High-Priority Phenotype and Exposure Measures for GWA Studies"). A third RFA ("Genome-wide Studies in Biorepositories with Electronic Medical Records") deals with GWAS data generation and infrastructure.

As for the second goal of improving analysis strategies, NHLBI has funded the ENDGAME project in which NHGRI has collaborated. A standardized database for genotypes and phenotypes, dbGaP has been created by NCBI (www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gap).

Several efforts have been taken to build successful collaborations across NIH, including: a multi-IC (Institute/Center) symposium in June 2006 on population resources and genomic technologies, the GAIN and GEI collaborations, discussions about the harmonization of data sharing policies across NIH, and pursuit of population genomics opportunities with other ICs. Key recommendations from the symposium on which progress has been made include making key elements of consent for GWAS available for IC staff, posting examples of successful consortium agreements, and defining database terms for GWA studies in dbGaP. Other workshops were held to produce rigorous algorithms to define approaches to follow-up GWA signals with sequencing and for establishing standards for defining the validity and replication of GWA findings. Another multi-IC symposium is scheduled for late May 2007.

Dr. Manolio then briefly described a set of actual and projected budget figures for population genomics. She noted that the only population genomics program funded in FY06 was a grant that was submitted in response to the RFA on public consultation. The largest component in the FY07 budget was the GEI initiative (which was funded by a special Secretary's Initiative in the FY07 appropriation, not by NHGRI). Awards under two NHGRI RFAs, on biorepositories and on phenotyping (see above) will be made in FY07, while the proposed Rapid Investigation

initiative would be funded in FY08. Projections for FY09 include programs on the population genetics of coumadin response, good health/disease resistance, and adverse drug reactions; work on clinical applications is also targeted for FY09.

Council asked about model consent language across the institutes and wondered what the ASHG may have developed. Dr. Manolio clarified that the goal was not to develop model consent language across institutes, but to collect individual consent forms from the institutes as models. Council also asked if there are parallel efforts with other federal groups that do population studies. Dr. Manolio told Council that her group has worked most closely with the CDC Office of Public Health Genomics and have also made efforts to work with the National Center for Health Statistics.

CONCEPT CLEARANCE: Rapid Investigation of Genetic Associations in Population Studies

Dr. Teri Manolio presented a Concept for an RFA on Rapid Investigation of Genetic Associations in Population Studies for Council clearance. The proposed initiative is to support the investigation, in well-characterized population samples, of genetic variants identified as potentially causally associated with complex diseases. The goal is to utilize existing cohort studies and clinical trials to determine the population impact of putative risk variants, identify modifiers of gene-trait associations, and to identify potential clues to gene function. Previous failure to pursue population-based investigations of important variants is likely the combined result of lack of appreciation of opportunity, lack of information on available population studies and included data, and lack of familiarity with the intricacies of databases and requisite analytic methods. NHGRI efforts to date have tend ed to focus on the lack of access to data, but hopes to address lack of information on available studies and lack of familiarity with analysis approaches with the Rapid Investigation initiative.

The design of the Rapid Investigation initiative is to support 3-5 investigative groups that work together to develop methods and procedures for, and then perform, genotyping and association analyses in population-based studies to characterize important genetic variants on a population basis. A Steering Committee would be formed for the identification of important variants and to facilitate the dissemination of results to the research community. A Coordinating Center would also be funded to handle the new data generated from the individual investigations.

Characteristics of suitable studies will include well-characterized, population-based samples of participants; high-quality DNA that has been or could be isolated; and populations of sufficient size and breadth to allow assessment of the wide range of variants and traits with sufficient power, including the broad range of participants approximating diversity of the U.S. population. Studies involving under-represented racial/ethnic minorities, particularly those with disproportionate burdens of disease, would particularly be sought.

The projects are anticipated to be funded as Cooperative Agreements (U01 mechanism) with $32M over a four-year period for 3-5 investigative groups and a Coordinating Center. Roughly one third of the costs would go towards genotyping and DNA amplification or isolation as needed in ~40,000 participants.

Council asked about the distinguishing characteristics of this program, and Dr. Manolio replied that the initial genome-wide association studies that were funded for GAIN and GEI are to discover the variants; these will be followed by replication and other steps to find a variant. The Rapid Investigation project would come after a strong candidate gene is found and is intended to allow description of the epidemiologic architecture, including population prevalence, associated risk, risk in population subgroups, and potential risk modifiers. Dr. Collins added that the intent is to go to the next level needed to understand the architecture critical for public health.

Council asked about the mechanisms for community engagement and education. Dr. Manolio explained that this program, as it strives to understand variants better, is a step towards communicating to the community. Council also questioned whether this was the right time for this program or if we should wait until more GWA studies are completed. Dr. Manolio noted that there already are 6-12 well-documented examples of putative causal variants that could be studied in this way, and there likely will be more very soon. Council suggested renaming the concept to "Coordinated Cohort Evaluation of Population Risk." When asked about incentives to encourage the reuse of cohorts, Dr. Manolio explained that researchers would be funded for the cohorts they applied for, not the number of genes that can be put through the study.

Council approved the concept clearance.

OTHER PROJECT UPDATES

Molecular Libraries

Dr. Carson Loomis discussed the Molecular Libraries Screening Center Network (MLSCN), part of the Molecular Libraries initiative, which is one of the 27 initiatives being undertaken by the NIH Roadmap. The main goal of the Molecular Libraries program is to generate molecular probes that will be useful research tools for the scientific community. The program is now two years into the pilot phase, during which time the Screening Centers have been started and begun development.

The Molecular Libraries program comprises several components, including technology development, screening, and informatics. The MLSCN (that includes nine pilot screening centers supported by cooperative agreements and one intramural center, the NCGC) is the hub of the program, around which the additional components (Chemical Diversity, Assay Development, Instrumentation, Predictive ADMET, Compound Repository, PubChem, and Cheminformatics Research Centers) revolve. The MLSCN utilizes compounds from the Molecular Libraries Small Molecule Repository (MLSMR) and all data generated are deposited into PubChem, a new NCBI public database. The MLSMR compound collection currently consists of 144,653 compounds. To date, 125 assays have been assigned to the screening centers and assay demand is increasing.

A Midcourse Review occurred in December 2006, from which the program received encouragement to transition to a production phase and gained valuable feedback on the

progress, goals, and directions of the program. One of the key issues for the midcourse review was an assessment of the productivity metrics of the first 18 months; most importantly, seven new probes have already been produced. During the remainder of the pilot phase, efforts will be focused on the development of the far end of the probe production pipeline, as the Centers build their chemistry capabilities to produce chemical probes of appropriate sensitivity and specificity.

During the pilot phase, the ten centers were funded. The structure for the MLSCN will change for the production phase, which will consist of 3-4 Comprehensive Screening Centers focused on high throughput probe production, 2-3 Specialized Screening Centers which will handle fewer, lower throughput high-content assays; and 2-3 Specialized Chemistry Centers which will provide additional chemistry capability for the Network, A solicitation for X02 pre-applications was released in April 2007, and applicants for the U54 centers will be invited in October. Funding is scheduled for July 2008. The total available for the production phase will be about $70 million/year for six years. A midcourse review will take place in 2010, and the Screening Centers will begin to transition out of Roadmap in 2011-2012.

Council asked what the MLSCN would consider as a definition of success. Dr. Loomis acknowledged that throughout the review process, the focus was on the number of assays being screened, but now that the program is 18 months old, publications and probes produced in Molecular Libraries are a focus. Publications and probes will be the measures of success in the long run.

Council found the goals to have an academic flavor and suggested that the program make more targeted goals as well; for example, having at least one high affinity probe for protein kinases of a certain type. When Council suggested picking categories for probe goals, Dr. Collins noted that the RICC (Roadmap Implementation Coordinating Committee) is currently having a discussion with the program staff about whether or not the new centers should have thematic components. When asked how assays are prioritized, Dr. Guyer explained that the management of assay prioritization involves a Project Team composed of one or more representatives from at least 12 ICs, so that the needs of a wide variety of NIH ICs are met by the assay selection.

Pathogens and Vectors

Dr. Lu Wang reported on the Eukaryotic Pathogens and Disease Vectors Sequencing Workshop, held by NHGRI and NIAID in November 2006 and co-chaired by Bill Gelbart and Claire Fraser-Liggett. One of the goals of the workshop was to help identify the most significant human eukaryotic pathogens and invertebrate vectors of infectious diseases for potential genomic sequencing as a collaborative and coordinated effort between NHGRI and NAIAD. A second goal of the workshop was to evaluate the usefulness and limitations of pathogen and vector genomic data in the study of infectious disease. Among the workshop's conclusions reached was the recognition that pathogen and vector sequencing projects have already provided genomic data sets that have begun to deliver significant benefits, but that the utility of this approach has not been exhausted. The workshop also concluded that there is a need for collaborative, coordinated efforts between NHGRI and NIAID.

The workshop attendees recommended that the highest priority should be given to *Anopheles*, *Plasmodium*, and *Trypanosome* species, based on the magnitude of the global health problems caused by these organisms. Also, they recommended sequencing additional genomes of eukaryotic pathogens and disease vectors, including additional isolates where a reference genome is available. Support for sequencing projects that enable comparative studies of pathogen and vector genomics, as well as studies of microbial population genomics, was recommended. Workshop participants also noted the value of facilitating genomic research on pathogens and vectors by funding the development of additional genomic resource for pathogens and vectors, not just the sequence; supporting unrestricted, rapid release of genomic data into international databases; supporting small-scale genome projects that would be informative for developing and committing to later-scale genome sequencing projects; and supporting the development of reagents and resources that will enable the use of genomic data. The workshop participants also recommended that an external scientific group of experts be established to provide guidance.

After the workshop, NHGRI and NIAID formed the Eukaryotic Pathogens and Disease Vectors Sequencing Target Selection Working Group to develop these high priority recommendations for sequencing targets and to identify community experts to help develop plans for the sequencing projects. Through these experts, several white papers are well under way.

Dr. Collins noted that this newly formed Working Group will be another means of ensuring that the sequencing pipeline will be used to its best advantage. He also mentioned that these sequencing efforts will follow NHGRI's policy on data release.

Council suggested adding to the working group people with expertise in the statistical and population genetics fields, and supported the recommendation to sequence host organisms.

STATEMENT OF UNDERSTANDING BETWEEN NACHGR AND NHGRI

Cheryl Chick presented a newly modified Statement of Understanding between NACHGR and NHGRI staff which serves as the agreement for review of grant applications by Council and NHGRI staff administrative authorities. The changes since May 2006 stem from the NIH reauthorization, which most notably requires Council review of all applications, even those with direct costs of $50,000 or less. Another modification was designed to ensure that critical Council operations will continue in the event of an emergency.

The Council adopted the statement of understanding.

COUNCIL-INITIATED DISCUSSION

Council suggested that a presentation on the TCGA project, addressing the acquisition of tissues as well as the technologies being used for analysis, take place at the September Council meeting. Staff suggested inviting Ann Barker from NCI who has led the effort, as well as potentially Matt Meyerson and Rick Wilson.

In addition, Council requested receiving feedback on the large cohorts that are available for large studies. The timing looks right to look at the NIH inventory of such cohorts in some detail, and so at some stage, some feedback would be helpful. Dr. Teri Manolio can collect information on this.

Council would also like an update, based on Dr. Eric Green's presentation, to follow-up on the 18% of NHGRI DIR funds being pumped into the NIH Central Management Fund and Clinical Center tap and strategies for utilizing these funds in the bench-to-bedside pipeline.

ANNOUNCEMENTS AND ITEMS OF INTEREST

Dr. Guyer directed Council to the Council folders containing items of interest.

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 208 applications, requesting $93,976,274. The applications included 37 regular research grants, 10 pilot projects, 12 ELSI grants, 114 RFA grants, 4 area grants, 3 center grants, 3 conference grants, 1 training grant, 1 continuing education training program grant, 11 SBIR Phase I grants, 9 SBIR Phase II grants, 2 fellowships and 1 STTR phase 1 grant. A total of 122 applications totaling $71,505,832 were recommended.

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

Signed 9/15/2007

_____ _____

Date

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

Signed 9/19/2007

_____ _____

Date

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