

**NATIONAL ADVISORY COUNCIL FOR HUMAN GENOME RESEARCH
SUMMARY OF MEETING¹**

February 9, 2009

The Open Session of the National Advisory Council for Human Genome Research was convened for its fifty-fifth meeting at 8:32 A.M. on February 9, 2009 at the Fishers Lane Conference Center, Rockville, MD. Alan Guttmacher, Acting Director of the National Human Genome Research Institute, called the meeting to order.

The meeting was open to the public from 8:32 A.M. until 3:42 P.M. on February 9, 2009. In accordance with the provisions of Public law 92-463, the meeting was closed to the public from 3:42 P.M. on February 9, 2009 until adjournment for the review, discussion, and evaluation of grant applications.

Council members present:

Michael Boehnke, *ad hoc*
Eric Boerwinkle
Mark Chee, *ad hoc*
Rex Chisholm, *ad hoc*
Richard Cooper, *ad hoc*
Jorge Contreras Jr.
Claire Fraser-Liggett
Richard Gibbs
Geoffrey Ginsburg
Caryn Lerman
Patrice Milos
Richard Myers, *ad hoc*
P. Pearl O'Rourke, *ad hoc*
Pilar Ossorio
David Page
Paul Sternberg Jr.
David Valle
Richard Weinshilboum

Council members absent:

None

Ex officio members absent:

None

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
Maggie Bartlett, OD
Tsegahiwot Belachew, DER
Vivien Bonazzi, DER
Vence Bonham, OD
Joy Boyer, DER
Courtney Bracey, OD
Lisa Brooks, DER
Comfort Browne, DER
Debbie Chen, DER
Cheryl Chick, DER
Monika Christman, DER
Christine Cutillo, DER
Camilla Day, DER
Karen DeLeon, OD
Laura Dillon, DER
W. Greg Feero, OD
Elise Feingold, DER
Adam Felsenfeld, DER
Colin Fletcher, DER
Phyllis Frosst, OD
Susan Garges, DER
Peter Good, DER
Bettie Graham, DER
Eric Green, DIR
Alan Guttmacher, OD
Mark Guyer, DER
Sarah Harding, OD
Lucia Hindorff, DER
Ephraim Johnson, DER
Christopher Juenger, DER

Heather Junkins, DER
Carson Loomis, DER
Teri Manolio, DER
Jean McEwen, DER
Keith McKenney, DER
Lisa McNeil, DER
Anika Mirick, DER
Ken Nakamura, DER
Ken Ow, OD
Vivian Ota Wang, DER
Brad Ozenberger, DER
Betsy Parker, OD
Jane Peterson, DER
Rudy Pozzatti, DER
Erin Ramos, DER
Eddie Rivera, OD
Cristen Robinson, DER
Ellen Rolfes, OD
Laura Rodriguez, OD
Anna Rossoshek, DER
Jeff Schloss, DER
Geoff Spencer, OD
Jeff Struewing, OD
Gary Temple, DER
Elizabeth Thomson, DER
Susan Vasquez, OD
Lu Wang, DER
Christopher Wellington, DER
Kris Wetterstrand, DER
Rosann Wise, OD
J. Julia Zhang, DER

Others present for all or a portion of the meeting:

Andrea Beckel-Mitchener, National Institutes of Mental Health (NIH/NIMH)
Joann Boughman, American Society of Human Genetics
Andrea DeSanti, Social & Scientific Systems
Jane Hammond, Research Triangle Institute International
R. Rodney Howell, American College of Medical Genetics
Kathy Hudson, Johns Hopkins University
David Kaufman, Johns Hopkins University
Thomas Lehner, National Institutes of Mental Health (NIH/NIMH)
Bobbie Peterson, Mayo Clinic Advanced Genomics Technology Center
Linda Rodman, Office of the Director (NIH/OD)
Rhonda Schonberg, National Society of Genetic Counselors
Joan Scott, Johns Hopkins University
Wendy Uhlmann, National Society of Genetic Counselors

INTRODUCTION OF NEW MEMBERS AND STAFF, LIASONS AND GUESTS

Dr. Guyer noted that the new Council slate had been approved and six members were now full members. They were participating at this meeting as *ad hoc* Council Members: Michael Boehnke, Mark Chee, Rex Chisholm, Richard Cooper, Richard Myers, and P. Pearl O'Rourke.

Dr. Guyer introduced new NHGRI staff: Susan Garges, Program Director, DER.

Dr. Guyer welcomed members of the press and liaisons from professional societies: Rhonda Schonberg and Joann Boughman from the American Society of Human Genetics, R. Rodney Howell from the American College of Medical Genetics, and Wendy Uhlmann from the National Society of Genetic Counselors.

APPROVAL OF MINUTES

The minutes from the September 2008 Council meeting were approved as submitted.

FUTURE MEETING DATES

The following dates were proposed for future meetings: May 18-19, 2009; September 14-15, 2009; February 8-9, 2010; May 17-18, 2010; September 13-14, 2010; and February 7-8, 2011.

DIRECTOR'S REPORT

I. GENERAL ANNOUNCEMENTS

Applications Invited for NHGRI Director

NHGRI is accepting applications for NHGRI Director through April 16, 2009. For more information, see <http://www.genome.gov/27529636> or contact Regina Reiter at (301) 402-1130. Applicants must possess an M.D., Ph.D., or comparable degree in the health sciences field. Applicants should have senior-level experience and comprehensive scientific knowledge of

research programs in an area relating to genetics or molecular biology. Lastly, applicants should have expertise in policy and ELSI issues relating to genetic research.

Linda S. Birnbaum, Ph.D. Named New Director of the NIEHS

In January 2009, Dr. Raynard Kington, acting director of the NIH announced the appointment of Linda S. Birnbaum, Ph.D., D.A.B.T., A.T.S., as director of the National Institute of Environmental Health Sciences (NIEHS) and the National Toxicology Program (NTP). Dr. Birnbaum was previously a toxicology and microbiology researcher and most recently served as a senior advisor to the Environmental Protection Agency (EPA).

Presidential Early Career Award for Scientists and Engineers

Two NHGRI intramural investigators, Elliott Margulies and Daphne Bell, were selected for the Presidential Early Career Award for Scientists and Engineers.

Elected to the Institute of Medicine

Several NHGRI-related colleagues were elected to the Institute of Medicine, including Dr. David Page of Whitehead Institute and MIT, Dr. Harry Dietz of Johns Hopkins University School of Medicine, Dr. Raju Kucherlapati of Harvard Medical School, and Dr. Nancy Kass of Johns Hopkins University Bloomberg School of Public Health.

NHGRI Planning Process

In April 2008, NHGRI senior staff and a small number of external experts held a retreat to further define the content areas that should be addressed in the future. NHGRI hopes to initiate conversations within our community by releasing a series of "white papers" that will begin to address relevant topics. A number of white papers on various topics will be posted in the coming months; four are now available for comment regarding the following:

Applying Genomics to Clinical Problems – Diagnostics, Preventative Medicine, and Pharmacogenetics, submitted by: David Valle, M.D., and Teri Manolio, M.D., Ph.D.

Applying Genomics to Clinical Problems – Therapeutics, submitted by: Harry Dietz, M.D., and Christopher Austin, M.D.

A Vision for the Future of Genomics: Education and Community Engagement, submitted by: Vence Bonham, J.D., and Sharon Terry, M.A.

The Future of Genome Sequencing, submitted by: Mark Guyer, Ph.D., and Adam Felsenfeld, Ph.D.

Please note that these are not white papers in the standard sense of a specific proposal. Rather, they are sets of questions that NHGRI, with input from a few outside experts, has proposed as the most important topics to address. The first goal is to develop a set of more refined and useful questions. Once that set has been established, NHGRI will then solicit the scientific community for answers to those questions. The Phase I comment deadline for these white papers is February 27, 2009.

II. NHGRI – EXTRAMURAL PROGRAM

Funding Opportunities. Dr. Guttmacher reported to Council that NHGRI is involved in several new funding opportunities, which are listed on the genome.gov, cancer.gov, nihroadmap.nih.gov, and nih.gov websites.

Genome Technology. New sequencing technologies are driving the boom in sequencing. “Cancer Genes” were named the #3 Breakthrough of the Year and the “Sequencing Bonanza” was named #10 by *Science* for 2008.

Advanced DNA Sequencing Technology Program. NHGRI’s Advanced DNA Sequencing Technology Program is accelerating technology development to bring about routine whole-genome sequencing. Jeff Schloss authored two papers that appeared in *Nature Biotechnology* in October, 2008: 1) Review article entitled “The potential and challenges of nanopore sequencing” which summarizes current views of the development of nanopore sequencing, and 2) Commentary piece entitled “How to get genomes at one ten-thousandth the cost” which describes the goals and progress of the Advanced DNA Sequencing Technology Program. Outlines for these papers were initially developed at last year’s grantee meeting.

The next Advanced DNA Sequencing Technology Program Meeting for grantees is scheduled for March 31-April 1, 2009. The final day is open to others who may have solutions to problems or can help generate new ideas that bear chemical or physical approaches to enable new sequencing methods. For more information, please contact Jeff Schloss (schlossj@mail.nih.gov).

Cancer Genome Progress.

Science recognized several high impact papers in cancer genomics by naming “Cancer Genes” as the #3 Breakthrough of the Year. Major reports, including TCGA, have begun to appear, giving way to new avenues for diagnosis and treatment of glioblastoma, pancreatic cancer, lung cancer, and AML. The Tumor Sequencing Project (TSP) Consortium sequenced >600 genes in approximately 188 lung adenocarcinoma specimens and discovered >1,000 protein-altering mutations. Results revealed new tumor pathway architecture and provided insights for therapeutic targets for treatment.

Encode and modENCODE. The ENCODE data release policy was finalized in October 2008 and the ENCODE Analysis Workshop was held December 7-8, 2008. A joint meeting of the ENCODE and modENCODE Consortium will be held on March 25-27, 2009. The meeting will focus on data integration and identification of production bottlenecks. A marker paper describing the scope and plans of the modENCODE Consortium is currently under review.

ELSI Natural Selection Workshop. The ELSI Natural Selection Workshop was held on October 28, 2008 to help identify issues for investigators, peer reviewers, and science editors to consider, as well as future directions for genomic and ELSI research regarding natural selection. A report on the meeting is forthcoming.

Statistical Genetics Training. NIGMS and NHGRI staff have developed an implementation plan based on the recommendations of the workshop report. The membership of the

International Genetics Epidemiology Society was surveyed about training and research support. The survey is still ongoing and about 10% of its members have responded to the survey thus far. The summary so far is that 73% of members had NIH funding, most had NIH training support, and most PIs could mentor additional trainees if funds were available. Over 85% of the respondents reported that if their pre-doctoral training slots were not filled, it was due to lack of qualified US citizens.

NIGMS and NHGRI staff sent letters to ASHG and the International Genetic Epidemiology Society (IGES) inviting them to collaborate in implementing some of the workshop recommendations to enhance statistical genetics and genetic epidemiology research. NIGMS and NHGRI staff are also planning ways to bring the Center for Scientific Review into discussions with professional societies to ensure that methodological applications get an optimal review.

III. NHGRI – INTRAMURAL PROGRAM

NIH Scientists Discover Crucial Control in Long-Lasting Immunity. Dr. Ronald Germain (NIAID) and Dr. Pamela Schwartzberg (NHGRI) have identified a protein called SAP which facilitates relationships between T and B cells, enhancing long-term immunity after infection. Individuals who produce less SAP protein suffer from lethal infections from Epstein-Barr virus, which is otherwise rarely fatal. These SAP-deficient individuals also have higher occurrences of B-cell lymphomas. Results were published in *Nature* on October 9, 2008.

Anti-Cancer Drug Prevents, Reverses Cardiovascular Damage in Mouse Model of Premature Aging Disorder. NHGRI investigators led a team that discovered that the anti-cancer drug, tipifarnib, prevents and reverses cardiovascular damage in a mouse model of progeria, a rare disorder that causes dramatic human premature aging. Results were published in the *Proceedings of the National Academy of Sciences* on October 14, 2008.

Scientists Identify Gene Variant Involved in Isolated Cleft Lip. A trans-NIH team including NHGRI researchers discovered that approximately 20% of isolated cleft lip incidence may be due to a SNP within a gene involved in facial development. This variant resides in a highly conserved stretch of DNA that is nearly identical in the genomes of the twelve species examined.

Researchers Discover New Genetic Risk Factors Involved in Adult and Childhood Obesity. An international consortium identified six new genetic variants associated with body mass index: *TMEM18*, *KCTD15*, *GNPDA2*, *SH2B1*, *MTCH2* and *NEGR1*. The study also confirmed two previously identified genetic associations with BMI: *FTO* and *MC4R*. Results were published in *Nature Genetics* on January 9, 2009.

IV. ROADMAP PROGRAMS

Human Microbiome Project (HMP). The HMP Jumpstart Consortium, comprised of the sequencing centers at the Baylor College of Medicine, The Broad Institute, the J. Craig Venter Institute, and Washington University, has made considerable progress. The consortium has >500 bacterial genomes either completed or in its sequencing pipelines; this is more than half of the

goal of establishing a genome data resource with the sequences of 1000 bacteria over the next four years. The consortium has also begun sequencing clinical samples collected from five body sites per donor. The Data Analysis and Coordination Center has been funded to collect and coordinate the project's data. HMP also participated in the launch of the International Human Microbiome Consortium, which is comprised of nine countries and which will coordinate international efforts to ensure rapid release of molecular and clinical data.

Epigenomics. The Roadmap 1.5 Epigenomics initiative-funded four Epigenomic Mapping Centers, and a Data Coordination and Analysis Center, plus R01 grants for Technology Development, and Novel Epigenetic Marks in September 2008. Applications addressing the Epigenomics in Human Health and Disease component will be reviewed this Spring and funded in the Summer of 2009. A project meeting was held in November 2008 to plan the implementation of the Reference Epigenomic Maps. A meeting on Emerging Evidence for Epigenomics in Human Disease is scheduled for March 16-17, 2009 and a meeting on Exploring International Epigenomic Coordination meeting is scheduled for March 17-18, 2009. Efforts are being made to coordinate with the ENCODE project (i.e., through the use of common ES cell lines).

V. NHGRI OFFICE OF THE DIRECTOR

Population Genomics. The Dark Matter Workshop was held on February 2-3, 2009 to discuss potential sources of "missing heritability." NHGRI's GWA catalog has expanded to include more than 240 studies and over 1,100 SNPs. GENEVA also funded six new GWAS proposals examining glaucoma, lung disease, and prostate cancer among others.

Understanding the Role of Genomics in Racial and Ethnic Health Disparities Workshop. NHGRI hosted a roundtable discussion on Race, Ethnicity, and Genetics in 2004. In September 2008, NHGRI partnered with NCI and NCMHD in a forum to discuss the interplay of the environment, social determinants, and genomics in racial and ethnic health disparities. Forty-five external participants from around the country joined more than a dozen NIH representatives to brainstorm and to develop recommendations for the scientific community addressing racial and ethnic health disparities. A report on the meeting is forthcoming.

Family History Activities. NHGRI's Genomic Healthcare Branch and the Office of the Secretary of HHS worked with a broad group of public and private stakeholders in 2007 and 2008 to develop new standards for family history, which have now been adopted by the Health Information Technology Standards Panel (HITSP). With the help of the Office of the National Coordinator for Health Information Technology (ONCHIT), a broad consortium of federal agencies and private partners has created a new version of the U.S. Surgeon General's popular public family history tool "My Family Health Portrait." The tool was released on January 12, 2009 and is available at <https://familyhistory.hhs.gov/>. This new version of the tool is more comprehensive, compatible with a large number of electronic health record and personal health record systems, and freely downloadable by vendors.

NHGRI Staff to Edit Genomic Medicine Series in *New England Journal of Medicine*. Dr. Greg Feero, Dr. Alan Guttmacher, and Dr. Francis Collins are set to begin editing a series on

medical genomics which will appear in the *New England Journal of Medicine* in early 2010. Planned topics include: introduction to genomics, disease risk assessment, health disparities in the age of genomic medicine, Type 2 diabetes and obesity, genomics and the continuum of cancer care, genomics and eye disorder, infectious disease, new therapeutic approaches for Mendelian disorders, mental retardation and autism, cardiovascular disease and stroke, Type 1 diabetes and autoimmune diseases, pharmacogenetics and drug development, and various ethical, legal and social issues.

NHGRI Sponsored Darwin Day Programs. NHGRI is sponsoring a NIH symposium entitled "Darwin at 200: Evolution, Genomics, and Medicine" on February 12, 2009. Speakers will include Dr. Alan Guttmacher and Dr. Eric Green, plus celebrated researchers from across the country. NHGRI and the Smithsonian Institution's National Museum of Natural History will also co-sponsor a program for high school students from the Washington metropolitan region. The Darwin Day celebrations will include a morning of interactive events for the students, followed by an afternoon program consisting of presentations and panel discussions.

VI. NHGRI – POLICY

Appropriations Update – American Recovery and Reinvestment Act of 2009.

Congressional focus has been consumed by the consideration and negotiation of the President's stimulus package, the American Recovery and Reinvestment Act (ARRA). In terms of dollars for the NIH within the package, the final House number was \$3.5 billion, which was increased to \$10 billion in an amendment negotiated between Senators Specter and Harkin. \$9.2 billion of the funds are devoted to research funding, the remainder for facilities, and shared instrumentation and other capital equipment needs. While it is impossible to predict how this will fare when the bill is sent back to the House for final consideration, it is anticipated that there will be pressure to maintain the higher level of funding.

Fundamental to the provision of ARRA research funds as part of the overall economic stimulus is that these dollars are not going into baseline NIH funding, but will be used to support short-term (one-year or two-year) new grants targeted to specific scientific challenges, the expansion of on-going projects in need of an infusion of funds, and particular public health priorities. NHGRI, along with the other ICs, has submitted suggestions for high-priority needs and timely opportunities for such "Challenge" grants. NIH leadership continues to review and discuss the options put forward in anticipation of final approval of the stimulus by Congress. There are also funds within the bill dedicated to increase Comparative Effectiveness Research, and the current plan is for NIH to receive an additional \$400 million dollars for this purpose, although again, this number may change in the final bill.

With regard to FY09 funding, Congress did not pass the FY09 appropriations bills before the start of the 2009 fiscal year on October 1, 2008, but instead passed an omnibus Continuing Resolution to fund the government through March 6th, 2009 at FY08 levels. NHGRI's FY08 appropriation was \$486,779 million (representing a 0.1% increase from the FY 2007 level), and this is the level at which NHGRI will operate until there is an FY09 appropriation. Although debate on an FY09 omnibus appropriations bill was scheduled for last week, the House postponed taking up the measure until after the President's Day recess to maintain focus on the

President's stimulus package. This may result in legislators not being able to meet the March 6th deadline and consequently needing to pass another temporary measure in order to finish business on FY09 appropriations and send a final omnibus bill to the President. The President is not expected to present his FY10 budget to Congress until March or April 2009.

Council members discussed various ideas for the use of stimulus funds that the NHGRI might get. One suggestion was to provide one to two years of funding to senior post-docs in light of the increasing scarcity of positions at universities due to the budget constraints imposed by the economic downturn. Another suggestion was to use funds to provide continuing support for laboratory technicians and other research staff whose jobs might otherwise be eliminated due to loss of grants and/or other budgetary restrictions. A third suggestion was to use some of the funds towards capital investment program to buy machinery and other research equipment. Drs. Richard Gibbs and Paul Sternberg agreed to draft a letter from the Council to NIH Acting Director Raynard Kington recommending these uses of stimulus funds as NIH priorities.

Genetic Information Non-Discrimination Act (GINA) Update. The Genetic Information Non-Discrimination Act (H.R. 493, S. 358) was signed into law on May 21, 2008. This law provides a baseline of protection for all Americans against discrimination in health insurance or employment decisions on the basis of their genetic information. The provisions of the law go into effect beginning in May 2009 for health insurance companies and in November 2009 for employers. The Department of Labor, the Department of Health and Human Services, and the Equal Employment Opportunity Commission are developing regulations to implement the law.

PUBLIC CONSULTATION ON LARGE COHORT STUDY

Dr. Kathy Hudson, of the Genetics and Public Policy Center at Johns Hopkins University, presented the results of an NHGRI-funded survey of public opinions regarding a large genetic cohort study. This study was an outcome of discussions at NIH over the past few years about the possibility of implementing a large prospective observational study similar to the UK Biobank, with the objective of understanding both genetic and non-genetic contributions to complex diseases. In 2004, NHGRI drafted a study design that recommended the recruitment at least 500,000 participants nationwide to participate in such a study. The proposal also recommended a preliminary study to assess public attitudes and personal motivations to participate in the study, which could inform the design of a subsequent project of this type.

Dr. Hudson's group surveyed approximately 4,600 Americans from 2006-2008 via online surveys, 15 two-hour focus group sessions in five major cities, and five town hall discussions. Special groups of interest included Pennsylvania residents with environmental exposures from Three Mile Island, and smokers in Jackson, Mississippi. The population selected to participate in the study closely matched U.S. Census figures. Respondents were randomized to various study scenarios to determine how compensation and receipt of individual research results might affect the willingness to participate.

The results indicated that 84% of respondents supported such a study and 60% would participate. Returning individual research results and increasing the compensation from \$50 to \$200 increased the willingness of the respondents to participate in the study. Decreasing the study

burden, in terms of effort and time required to participate, did not significantly improve the willingness to participate. Returning individual research results was the most important factor to improve willingness to participate, as three of four respondents said that they would be less likely to participate if the results were not returned. Support of the study and willingness to participate varied little among demographic groups, although variations in influences (returning research results, increasing compensation, and decreasing study burden) were observed. Many respondents also wanted a study contract and consequences for researchers if such contracts were violated. The researchers concluded that there is widespread support in the general public for a large, national cohort study. In order to increase participation, the group recommended provision of adequate compensation and of individual research results.

Council asked Dr. Hudson whether the participants were asked if they had health insurance. The study did include questions about health insurance and frequency of physician check-ups, but these two factors did not make a difference in research findings. However, respondents did perceive a health benefit from participating in the study. It was noted that correlating ethnicity with health insurance and willingness to participate would be an interesting follow-up study.

Council further inquired about the amount of detail provided to participants regarding the return of individual research results. Participants were given comprehensive information, although the amount of time describing the research results was limited. It was noted that examining the amount of detail provided to participants in the research report would be an interesting follow-up study.

Council next asked about the order in which information was presented to participants and whether that affected their willingness to participate. The researchers did vary the order in which information was delivered in order to determine what is most effective, and they hope to continue to perform more such experimental manipulations in future follow-up studies.

Council expressed concern that returning individual research results to study participants is costly, burdensome, and irrelevant for non-clinical studies. NIH does not account for this mechanism in project design or funding. Dr. Hudson responded that individuals were briefed that this study is not a healthcare application, rather for research purposes only. The Office of Population Genetics is currently examining the financial aspects of returning research results through several programs including GWAS, PAGE, EMERGE, etc. Dr. Hudson and colleagues hope to further engage participants in consideration of the trade-offs of returning results.

NEW PEER REVIEW PROCEDURES AT NIH

A trans-NIH Peer Review Oversight Committee (PROC) was formed in June 2007 to address four areas of peer review: 1) engaging the best reviewers; 2) improving the quality and transparency of review; 3) ensuring balanced and fair reviews across scientific fields and career stages, and reducing the administrative burden of the review process; and 4) continuing evaluation of the peer review process. The PROC held internal and external town meetings, conducted surveys, and solicited feedback from investigators, and developed the following recommendations.

With regard to engaging the best reviewers: 1) allow continuous submission of applications from standing study section members; 2) allow review service to be spread over six years to provide more flexibility to reviewers; 3) pilot testing of "virtual reviews" using high bandwidth technologies to eliminate reviewer travel; and 4) development of a toolkit of ideas and approaches (incentives) to recruit the highest quality reviewers.

To address the quality and transparency of review: 1) implement a new scoring system with an impact score ranging from 1 to 9; 2) obtaining scores from assigned reviewers for each of five core review criteria for all applications (including streamlined applications); 3) revise (enhance) the review criteria, placing more emphasis on the significance/impact of the proposed research; 4) introduce new templates for written critiques, using different templates for different mechanisms; 5) shorten the application, with a 12-page limit for R01s, and with other mechanisms shortened appropriately; and 6) revising sections of the 398/424 application form to align with the revised review criteria (i.e., add a new section on the significance or impact of the proposed research).

To address the balance and fairness of reviews across scientific fields and career stages, and reduce administrative burden: 1) adopt an NIH policy to fund meritorious science earlier; 2) enhance success rates of new and resubmitted (A1) applications by eliminating the A2 submission; and 3) review similar applications together by identifying Early Stage Investigators (ESI) and New Investigators and clustering them during peer review process. In 2009, where possible, NIH will cluster new investigator applications (including ESIs) for review. The same approach is being considered for clinical research applications.

To address the continuous review of peer review: 1) establish an evaluation group to develop methods by which continuous review of peer review activities can be done; and 2) do some baseline surveys of NIH staff and the extramural community – these will be rolled out in March.

The following is a timeline for these major changes:

- Applications submitted for January/February, 2009 receipt dates for October 2009 Council round will not be allowed an A2.
- The new scoring system, enhanced review criteria, and critique templates will all be used for January 2009 submissions in the summer 2009 review meetings, and October 2009 Council round.
- The reduced page limits (12-page R01s) and revised section heads for 424/398 applications will be implemented by the January/February 2010 submissions, for summer 2010 review meetings, and October 2010 Council round.

One major effect that these changes will have is that there will be more "ties" among the scores, so funding decisions will have to involve other factors such as research priorities and portfolio balance. In this regard, the shorter application is intended to encourage the reviewers, program staff, and Council to focus on the significance of the project and decrease the attention paid to the details of the proposed methodologies. Finally, program staff and Council should be aware of the possibility that the shortened applications may incur a disadvantage to new investigators who have not had the opportunity to establish a track record of accomplishments and productivity.

Council noted that changes in the scoring system will increase the number of applications for review in future Council rounds. It is hoped, however, that changing the scoring system will eliminate "scoring compression" and increase the accuracy and precision in evaluating applications. Current and future changes to the peer review system can be found online at <http://enhancing-peer-review.nih.gov>.

NHGRI PROGRAM PORTFOLIO

Dr. Guyer presented an overview of the NHGRI program portfolio, as Council had requested at the September 2008 meeting. The information he described included: the various award mechanisms used by NHGRI, the total number of awards of each type, and the total amount dollars spent on each. The major NHGRI award mechanisms, in terms of both numbers and dollars, are U54s, U01s, R01s, P50s, and P41. Dr. Guyer also provided a breakdown of solicited versus unsolicited grant types, noting that the majority of NHGRI funding goes for programmed research, and information on the number of awards to new investigator awards. He then broke down the extramural spending pattern for each program, including sequencing, informatics, and functional genomics sequencing technology, CEGS, human genetics, population genomics, and ELSI.

The level of funding of the sequencing program decreased slightly from 2007 to 2008 and the funds were transferred to population genomics and non-sequencing technology development. It appears that the program can be maintained at this slightly lower funding level, since throughput has substantially increased. There was discussion about whether the sequencing program budget can be reduced further or whether it should remain at the same level. This issue will be revisited at the next Council meeting.

The FY09 budget is currently unclear, as NIH is operating on a Continuing Resolution that extends until March 6. The final FY2009 appropriation is still to be determined. Unlike other institutes, NHGRI does not fund according to a strict pay-line. The institute is small enough that applications can be individually considered each round. However, the "effective" pay-line is a priority score of 190 +/- 20-30.

BIENNIAL REPORT ON INCLUSION OF WOMEN AND MINORITIES IN NHGRI-SUPPORTED RESEARCH

Ms. Joy Boyer presented the biennial report on the inclusion of women and minorities in NHGRI clinical studies. This report is required by the NIH Revitalization Act of 1993 and is intended to help ensure that NHGRI is compliant with the NIH mandate to include women and minorities in clinical research. The data presented cover extramural and intramural research reported in FY 2007-2008. All human subjects research must report gender and minority inclusion unless exempted. Data are provided by investigators as "target data" in grant/study applications and then as "actual enrollment" data in annual progress reports.

The FY 2007-2008 NHGRI Report includes 1) FY 2007-2008 Extramural – both Actual Enrollment and 2008 Office of Population Genomics (OPG) Target Data, 2) FY 2007 – 2008

Intramural Actual Enrollment Data, and 3) a comparison with the 2007 NIH Aggregate Data and 2000 Census Data.

NHGRI Extramural Research: The FY 2007 Extramural Research Actual Enrollment (n=2,349) and the FY 2008 Extramural Research Actual Enrollment (n=5,566) data were presented. In both years, the majority of participants were white (~70%), non-Hispanic (~78%) and female (~63%). Approximately 15% did not provide a racial or ethnic identification.

OPG Target Data: The FY 2008 Office of Population Genomics Target Enrollment (n=222,433) data was presented. The majority of participants were white (~73%), non-Hispanic (~84%) and female (~66%). OPG enrollment data are not considered "Actual" until genotyping is successfully completed and data are included in databases. However, it was noted that OPG Target data accurately reflects Actual Enrollment. OPG studies' target enrollment dwarfs current DER actual enrollment data (n=222,433 in comparison to 5,566) and will largely determine future NHGRI actual enrollment counts.

NHGRI Intramural Research: The FY 2007 Intramural Research Actual Enrollment (n=79,023) and the FY 2008 Intramural Research Actual Enrollment (n=69,943) graphs were presented. In both years, a majority of participants were white (~61%), non-Hispanic (~70%) and female (~48%). Approximately 20-25% provided no racial, ethnic or gender identification.

The main conclusions of the presentation were that 1) NHGRI demographic data are comparable to NIH aggregate data, and 2) with the inclusion of the OPG data, NHGRI data will likely exceed the census levels for demographic diversity.

Council asked how human subjects research was defined. Ms. Boyer replied that each IRB has the responsibility for interpreting the definition human subjects research and the interpretations vary somewhat among institutions. However, regardless of how IRBs coded their studies, this report included all of the GWAS data to remain consistent with the spirit of the count. It was also noted that the self-reporting ethnicity parameter is difficult to capture since many Hispanic persons do not distinguish between race and ethnicity.

INFORMATICS AND NEW SEQUENCING TECHNOLOGIES

Dr. Vivien Bonazzi gave a presentation on the informatics challenges of the new sequencing technologies. The main factor to be considered is that the next-generation sequencers will generate terabytes of data, with estimates of ~250-500 terabytes in 2009, ~500-1,000 terabytes in 2010, and ~1-2 petabytes in 2011.

Infrastructure issues include: data storage (e.g., intensity files, raw and processed data, enhancement of current repositories and creation of new ones, large capacity required, types of storage, distributed storage, data compression, security, etc.), data transfer rates (e.g., bandwidth between computing resource and user, distributed storage, speed of access between various storage disks), data representation (e.g., standardized formats to simplify data exchange, data indexing for faster ways to traverse huge raw datasets, and metadata to include descriptors of

data generation), and computing capacity (e.g., increased CPU capacity, distributed CPUs or cloud computing, large-scale CPU clusters, security and bandwidth issues, etc.).

Analysis needs include: new tools for data analysis, the development of new analysis tools, scale-up of current tools to handle larger data volumes, data visualization, and statistical analysis.

NHGRI is planning to hold a Workshop in May or June 2009 to discuss how the Institute should approach these infrastructure and analysis needs. Planned participants in the workshop include community experts who deal with large sequence datasets (i.e., Genome Sequencing Centers, NCBI, EBI, UCSC, Bioconductor, other NIH Institutes, etc.) and other organizations that handle vast datasets (i.e., NSF supercomputing centers, DOE, NASA, NSA, Google, Amazon, Pixar, Aspera, etc.). A small planning committee will ensure the meeting is balanced with representatives from the biology and computing communities.

Council drew attention to the growing amount of clinical data (i.e., image files) in addition to the growing amount of sequence data. Council recommended including bioinformatics storage issues in the letter to Dr. Kington regarding use of stimulus funds.

PROJECT UPDATES

Genotype-Tissue Expression (GTEx). Dr. Jeff Struewing presented the GTEx program, a recent Roadmap initiative that is led by NHGRI, NCI and NHLBI, and aims to create a new community resource consisting of correlated human gene expression and genetic variation data. The expectation is that this database will allow users to follow up the results of GWAS studies through analysis of the relationship between GWAS data, which examines genetic variants associated with clinical traits, and expression quantitative trait loci (eQTLs), which are genetic variants associated with gene expression traits. This should allow identification of candidates for further experiments to determine the casual basis of GWAS association, as well as providing further understanding of natural genetic variation as perturbations of gene regulation, both of which should shed light on the genetic factors within and outside of coding exons that contribute to common human diseases, such as heart disease, diabetes, cancer, etc.

GTEx has been approved as a two-year pilot project to demonstrate the feasibility of obtaining DNA and RNA from many tissues from 160 donors. The donors will be identified in post-mortem autopsy or organ transplant settings. Living surgery patients will also be included to allow some comparisons of results from living and post-mortem tissue. If the pilot phase is successful, the project hopes to expand to include over 1,000 donors.

NHLBI released a funding opportunity on December 11, 2008 seeking a small business with the capability to serve as a Laboratory, Data Analysis, and Coordinating Center (LDACC). NCBI is preparing to develop an analysis pipeline to combine GWAS and eQTLs data. NCI is taking the lead on tissue procurement efforts. Funding will begin in FY10 and organizers are still in the preliminary RFP writing stages.

Council discussed whether expression arrays or transcriptome sequencing using next-generation technologies would be better. The organizers originally planned that the pilot would use

expression arrays while testing next-generation sequencing approaches on a smaller number of samples, and then re-examine this question closer to FY10. As a result of Council discussion, the use of next-generation sequencing earlier in the pilot is being considered.

Mammalian Gene Collection (MGC). MGC is a trans-NIH project led by NCBI, NHGRI, and NCI, that aims to identify and sequence full-length open reading frame (cDNA) clones for human, mouse, and rat genes. The MGC concept proposal was first presented in 1999, with first results reported in 2002, and last update provided in 2004. In 2005, Genome Canada added bovine cDNAs to the project. In total, MGC sequenced and verified cDNAs for a non-redundant set of 17,432 human, 17,562 mouse, 6,399 rat genes obtained by various methods, including 5' expressed-sequence tags (EST)-based isolation of cDNA clones, PCR rescue, and DNA synthesis. Random EST-based sequence clones initially provided more than half of the genes for the mouse and human genomes. As that approach reached the point of saturation, the other approaches were implemented to reach completion. Results indicated that the success of the PCR-rescue approach was most successful for those genes with higher numbers of known ESTs. Likewise, DNA synthesis success increased with smaller target sizes. It was also determined that mouse genomes have a higher mutation rate (0.91 freq per kb) than human genomes (0.70 freq per kb).

Comparing MGC full-CDS clone availability with the RefSeq database, the human collection is 91% complete, the mouse collection is 89% complete, and the rat collection is 41% complete (the objective was never to assemble a complete collection of rat cDNAs, only a specified number which was obtained). There is high concordance among human and mouse full-CDS clone availability with high-confidence RefSeq entries (defined as 10+PubMed results) at respective levels of 95% and 93%, reflecting upon the success of the project.

All MGC sequences are deposited in GenBank and are available without restriction. The cDNA clones generated by MGC are available through the IMAGE clone distribution network and are accessible to the community.

Council discussed how the inclusion of splice isoforms in a related project could enhance the value of MGC. It was also noted that clones generated via DNA synthesis likely represented real transcripts since researchers looked for strong RNA and evolutionary support. However, it is possible these clones could be sequencing artifacts.

1000 Genomes. Dr. Lisa Brooks provided an update on the 1000 Genomes project, which is being carried out to provide an improved resource for supporting genome-wide association studies. The goal of the project is to find 95% of the variants that are present in the human genome with a frequency of 1% (<0.5% in gene regions), and place them on haplotypes. The project expects to discover structural variations and SNPs alike, and for the data set to be robust enough for applications in various populations. In principle, this resource would contain almost all variants in any region of a GWAS hit, thus providing the set of almost all possible causal variants, so the GWAS investigator(s) would not need to sequence their samples, but could proceed in a more informed way to functional studies.

The 1000 Genomes project involves nine sequencing centers: Broad Institute, Sanger Institute, Baylor College of Medicine, Washington University in St. Louis, Beijing Genomics Institute, Applied Biosystems, Illumina, Max Planck (Berlin), and Roche. NCBI and EBI serve as the Data Coordinating Center. The latest data freeze (number four) included 2,815 Gb of sequence from low-coverage samples, 962 Gb of sequence from trio samples, and 61 Gb of sequence from gene regions. In total, 3,838 Gb of data have been produced. The fifth data freeze is expected to occur in the near future. There have been two data sets released to date, in December 2008 and February 2009. It has been determined that next-generation sequencing technologies have a 93% SNP detection rate for alleles at 4X coverage and 97% SNP detection rate for alleles at 5X coverage. Over 50% of discovered SNPs are novel. The project intends to include another 400-500 samples from across the world during the scale-up phase.

PhenX (Phenotypes and eXposures). Dr. Erin Ramos presented the PhenX project, a three-year project intended to identify standard phenotype measures (15 standard measures for each of the 20 high priority research domains), and to disseminate these measures to the research community through the PhenX Toolkit to facilitate future cross-study analysis.

In consultation with domain experts, 20 priority research domains pertinent to public health and genetic research were identified: alcohol, tobacco, and other substances; anthropometrics; cancer; cardiovascular disease; demographics; diabetes; environmental exposures; gastrointestinal; infection disease and immunity; lung function; neurological; nutrition and dietary supplements; ocular; oral health; physical fitness and physical activity; psychiatric; psychosocial; renal function; reproduction; and skin, bone, muscle and joint. Working groups were formed around each research domain and scientific experts in each field worked to select 15 high priority measures and standardized approaches to measurement.

These measures are available through the web-based PhenX Toolkit at www.phenx.org, which had a "soft release" on February 6, 2009. Researchers can use the Toolkit to add standard measures to ongoing studies, to consider PhenX standards when planning new studies, and to obtain high quality measures outside of their area of experience. By using the PhenX standards, researchers can ensure that their study will be compatible with others that also incorporate them, and combine studies to increase statistical power and improve their ability to identify genes associated with complex diseases.

Council noted the importance of seeking wide-spread endorsement and centralization of PhenX. It was suggested to increase communications between PhenX, caBIG, and dbGaP. There was discussion of requiring funded investigators to use PhenX measures, similar to the way caBIG requires funded investigators to use their Clinical Data Elements. Dr. Ramos noted that the Working Groups contained many IC liaisons and she will continue to work on centralizing the PhenX Toolkit.

COUNCIL-INITIATED DISCUSSION

Agenda items for the May 2009 Council include a presentation of the NHGRI Intramural Program by Eric Green and Clinical Program by Bill Gall, updates on sequencing strategies, and a report from the International Data Release Meeting, among other topics.

