Report of the NHGRI Training and Education Task Force, October 22, 2019

Task Force members included: Bruce Birren, Carol Bult, Jeff Struewing (co-chair), Jonathan Pritchard, and Wendy Chung (co-chair)

NHGRI's training portfolio (principally F-, K-, T-, and R25 grants) accounts for approximately 4.2% of the Institute's 2018 extramural budget (**Figure 1**). Despite significant training needs in the field of genomics, as delineated in this report, this proportion is well below most other NIH Institutes and Centers (**Figure 2**). The NHGRI Training and Education Task Force was formed to evaluate the Institute's current extramural training efforts and to identify potential growth areas. To develop this report, over the course of 4 months, the Task Force: (1) reviewed NHGRI's current portfolio in terms of the types of grants that have been awarded (T32s, Ks, Rs, etc.) and research topic areas (e.g., basic genomics, genomic medicine, and ELSI); (2) reviewed training-related input from strategic planning meetings; (3) received input from the Intersociety Coordinating Committee for Practitioner Education in Genomics (ISCC); (4) evaluated responses from approximately 30 colleagues from academic extramural research and industry environments to a set of questions that we posed to them; and (5) received feedback from council members at the September 2019 National Advisory Council for Human Genome Research meeting.

Table 1 delineates the landscape of education and training gaps, objectives, and strategies for various career stages. The needs are vast but not all stages and strategies are appropriate for significant grant-related funding by NHGRI. For example, the K-12 student group is large, with great needs and potential benefits, but it may well be outside the scope of NHGRI to support through significant grant funding. This target group and others may be better addressed by the NHGRI Education and Community Involvement Branch.

In addition to the specific groups and strategies identified in Table 1, the Task Force identified five core principles for training in genomics and data interpretation and in ELSI issues that will impact all target groups. These principles are as follows:

Diversity/Inclusivity – For genomics and genomic medicine to have a broad and equitable impact on our society it is paramount that training programs be accessible to individuals of diverse ethnic/racial backgrounds and socioeconomic statuses, sexual and gender minorities, and the disabled. It is important to train investigators on how to sensitively deal with the data from the communities that have historically been underrepresented and marginalized.

For postdoc and early faculty levels in computational biology, functional genomics, and related disciplines, women scientists are and should continue to be considered an under-represented minority (URM). Long-term goals should include the development of a workforce that includes clinical, research, and ELSI investigators that reflects the diversity of our nation's population and a genomically literate lay public. NHGRI is encouraged to contribute to NIH-wide or other efforts that continue to collect data on the diversity of trainees/workforce and ensure these findings are incorporated into training programs to better utilize information about the factors

that lead to the historic exclusion of certain groups in education and training and from the workforce. Individual and institutional training programs should be encouraged to document their use of data-driven approaches to recruitment, support and retention of trainees, particularly for those from historically underrepresented groups, including mentor-training for research advisors. We also need to train investigators on how to sensitively work with communities, and with the data from communities, that have historically been underrepresented and marginalized.

Multi-disciplinary - Genomics is an inherently multi-disciplinary field requiring expertise in diverse disciplines such as biology, biochemistry, engineering, physics, mathematics, bioinformatics and computer science. The application of genomics in medicine, likewise, requires different perspectives and expertise from technology, analytics, ELSI issues, and patient care. Within the area of patient care, there is diversity including genetics professionals, physicians, physician assistants, nurse practitioners, and nurses. Training initiatives should emphasize how to develop and manage such collaborations in an effective and ethical manner.

Modular and Just-in-Time Training Curricula – The rapidly changing nature of genomics technology and the application of genomics to medical practice demands new approaches to training in addition to traditional programs. We need options to educate new trainees but also established investigators and the medical workforce as the field evolves. The most effective way to meet these vast educational challenges is by developing curricula and educational modules that can be expanded as the field continues to evolve. Education modules that are components of curricula tailored to specific user groups should be designed to adapt rapidly to advances in genomics and data science technologies and be available "on demand" and, where appropriate, integrated into the clinical workflow. We expect for there to be ongoing need to update these modules and to do so rapidly to keep the resources current and maximally effective. As these modules are built, it will also be important to develop effective dissemination plans to ensure they are maximally utilized.

Data and Knowledge Democratization – Access to data and platforms for analysis and interpretation of genome data remain key challenges in genomics. Training initiatives should include grounding in state-of-the-art approaches to semantic and technical data standards that are fundamental to the democratization of data access and to supporting data reuse.

Foundational Knowledge – Given the speed and volume at which genomic data are generated, it will be important for those seeking advanced degrees to be able to manipulate these data. Acquisition of these foundational quantitative analytical skills begins in high school. There will be a need for those who: (1) have thorough knowledge in bioinformatics, computational biology, and statistics so that they can develop methods and analytic procedures to manipulate large data sets and (2) use the data generated from experiments to have a working knowledge of these disciplines so that they can interpret the data. These needs should be addressed appropriately in all training programs. To increase the diversity of genomic investigators, there should be efforts for focused support (postbac, bootcamps or summer programs) to improve

analytical and computational skills for motivated trainees immediately prior to or at the beginning of genomics training.

Given the ever-expanding impact of genomics on numerous scientific fields and in clinical medicine, the rapidity with which the field is changing, and the significant needs in terms of workforce training, the Task Force recommends that NHGRI significantly expand its training and education efforts. **An increase of 2% to a total of ~ 6% of the total extramural budget over the next 5-10 years would put NHGRI near the median of other ICs. The Task Force recommends an increase of 2-3%, which could be structured as a $6-$8M (1.5-2%) increase over 5 years, followed by a reevaluation and eventual increase of another $2-$4M (0.5-1%) within 10 years.** Even this large increase in funding cannot address all the needs identified in Table 1 and the core principles outlined above. Almost all areas identified in this report would benefit from increased efforts, but the Task Force identified the following areas as those that might benefit most and where NHGRI has the potential for the greatest impact from an increased investment in training and education.

Recommendations for NHGRI to consider in support of training and education initiatives

1. Increase funding for T32, F, and K grants and URM initiatives. Larger proportionate increases in funding should be focused on URM, K- and F- grants, followed by T32 grants. Historically, supporting trainees later in their careers who are dedicated to genomic research has proven more effective at developing a genomic research workforce. F awards are a mechanism to support genetic counselors who wish to pursue a PhD. New T32 grantees should be prioritized over increasing the size of existing programs, although new programs need to be of reasonable size for efficiency. A doubling of F-awards (currently <$500K/year), a 50% increase in DAPs and K-awards, and a 17% increase in T32s and the remaining training budget would result in a $5.5M increase.
2. Expand opportunities for individuals who need to round out their training to be effective genomic researchers (e.g. genetic counselors, ABMGG Laboratory Genetics and Genomics and other fellows, nurses and nurse practitioners, etc.). Such individuals have already demonstrated relevant interest and skills in genomics but are often not prepared to contribute as effectively as they could to research efforts because they have not been trained with all the relevant skills. In particular, many of these individuals could contribute significantly to ELSI and clinical implementation research.
3. Develop or expand efforts to increase the pool of data science and bioinformatics researchers, particularly among women and under-represented minority populations. Staff are encouraged to think creatively about training in this area and for #2 above, such as data science boot camps. focused on advanced analytical and programming skills development. Approximately $1M could be devoted to pilot programs in areas #2 & #3.
4. Expand or develop new award mechanisms to support the development and dissemination of training modules described above. Approximately $1M could be devoted to pilot programs for modular training development.

5. Set milestones and monitor progress toward the above goals. NHGRI staff should set ambitious goals and milestones both with regard to the number and size of awards, but also, to the degree possible, to increasing representation of women and under-represented minority populations. Staff are encouraged to include language in Funding Opportunity Announcements that set expectations about increasing under-represented minority populations. Individuals funded as trainees who go on to play key roles in team science genomics and data science initiatives or to industry but do not go on to become independent investigators in an academic setting should be considered successes.

The monetary targets outlined for 1-4 are approximate but would result in a $7.5M (43%) increase over the current training budget over 5 years, representing an additional 1.9% of the entire NHGRI extramural budget.

Table 1. Landscape of training and education gaps, objectives, and strategies for various career stages. The goal is to expand and support the pipeline of training for the next generation of traditional genomics trainees, add new genomic medicine researchers (physicians, genetic counselors, genetic diagnostic laboratorians), keep the research community up to date with advances in genomics, support integration of genomic medicine by the broad medical workforce, and educate the lay public.

Career stage	Gaps to fills	Objective of training	Suggestions for educational strategies
K-12	Genomic literacy	General understanding of genomics, increase interest in/awareness of careers related to genomics	Teacher plans and activities (extracting DNA from strawberries, discussion topics for the classroom-Grey's Anatomy episodes), recorded talks from experts
Undergraduate/ Postbac	Introduction to genomics	General understanding of genomics, increase interest in/awareness of careers related to genomics	Educational modules/laboratory activities
Graduate	Multiple aspects of genomics (data analysis methods, animal models, variant interpretation, medical applications, therapeutics, ELSI research methods, ethical aspects including working with industry, how to collaborate)	Provide up to date education on data analysis methods and new techniques, foster team science across disciplines and between academia/industry	Educational modules, short courses, more T32 and F training grants
Postdoc	Multiple aspects of genomics (data analysis methods, animal models, variant interpretation, medical applications, therapeutics, ELSI research methods, ethical aspects including working with industry, how to collaborate)	Provide up to date education on data analysis methods and new techniques, foster team science across disciplines and between academia/industry	Educational modules, short courses, more T32 and K training grants

Laboratory Genetics and Genomics (LGG) fellows*	Variant interpretation, additional research methods including data analysis methods, medical applications, ethics)	Teach robust methods of clinical genomic analysis and teach research methods for data analysis, foster team science	Educational modules, funding mechanisms for LGG fellows
Researchers	Data analysis methods, variant interpretation, medical applications, therapeutics, ELSI issues and methodology	Keep genomic researchers up to date with current methods as they are developed.	Online lectures, short courses at national meetings
Physicians	Variant interpretation, medical applications, ELSI issues	Educate physicians who are not geneticists how to use genetics in their specialty	Curriculum development, online modules, on demand education for common clinical scenarios, short courses at national meetings, cross training of fellows, certificate programs
Genetic counselors	Variant interpretation, medical applications, ELSI research methods	Teach robust methods of clinical genomic analysis and teach research methods for ELSI research to provide a new researcher pool	Modules for genetic counseling graduate curriculum, short courses at national meetings, funding mechanism genetic counselor research training fellowship
Other allied health professionals (nurses, social workers)	Basic variant interpretation, medical applications	Effective utilization of genetics in medical care	Online modules, on demand education for common clinical scenarios,
General public	Genomic literacy	Dispel myths about genomics and set appropriate expectations	TED style talks, Ask me Anything, social media engagement.

* Fellowship of the American Board of Medical Genetics and Genomics (ABMGG)

Figure 1. FY18 Training Awards

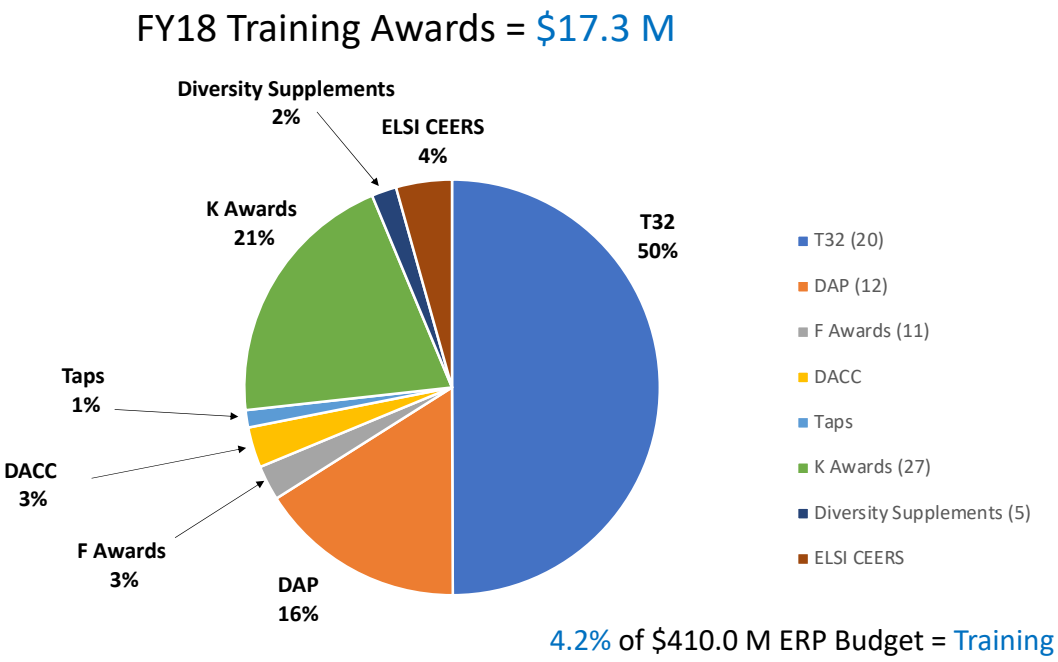
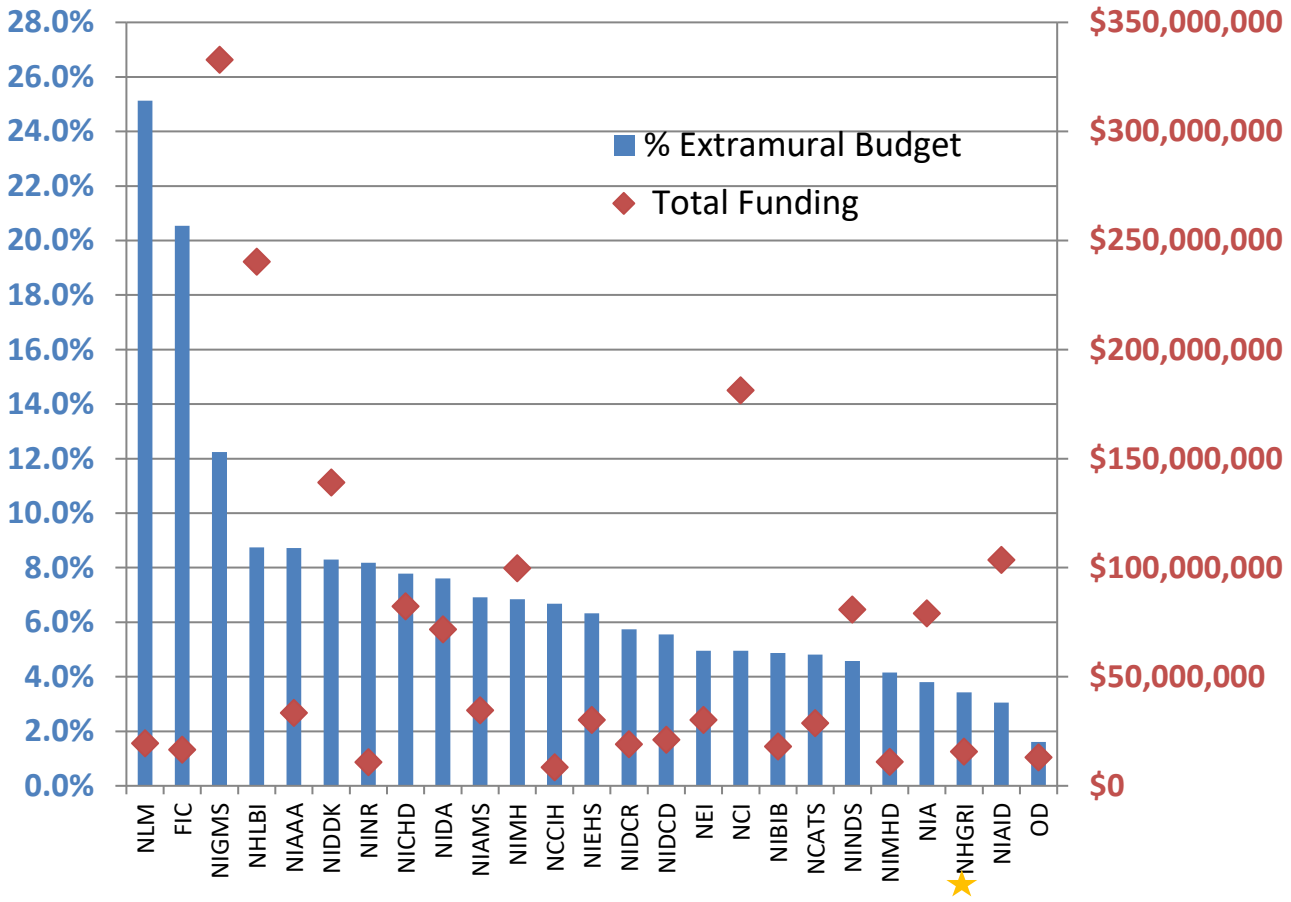


Figure 2. T-F-K-R25 Investments per IC (FY 2018)



Note – T-F-K-R25 represents only a subset of the total training investment in Figure 1, but a large subset that facilitates direct comparison to other NIH Institutes.