National Advisory Council for Human Genome Research

February 22, 2021

Concept Clearance for FOAs

Non-Human Primate Developmental Genotype-Tissue Expression (NHP dGTEx)

Purpose

The National Human Genome Research Institute (NHGRI) proposes a new initiative in non-human primates (NHP) to complement the recently launched Developmental Genotype-Tissue Expression (dGTEx) initiative. The goal of the dGTEx initiative is to establish a resource database and associated tissue bank to study gene expression patterns in multiple reference tissues during human developmental stages. NHGRI is now proposing a set of parallel non-human primate (NHP) tissue collections for comparative genomic assessment of tissue-specific gene expression at distinct developmental stages. This will involve the collection and RNA sequencing of multiple tissue samples from species representing Old World and New World primates at prenatal and postnatal timepoints. Comparative analysis of developmental expression profiles between primates and humans combined with existing data from other model organisms will help to determine common developmental networks and pathways as well as to identify developmental patterns that are primate- or human-specific. In addition to providing important knowledge about gene function and developmental evolution, this resource will also improve our understanding of similarities across species and inform selection of appropriate laboratory models for specific human developmental processes and stages.

Background

Many of the factors that influence health and optimal function throughout the lifespan are known to be initiated in prenatal and early childhood development. Through studies of developmental genetics, primarily in model organisms, we know that the expression of many genes is not only tissue-specific but also time-dependent and controlled through coordinated regulation of complex networks and pathways. Disruption of these processes can produce profound developmental errors and result in severe phenotypes. However, smaller variations in sequence and expression levels may influence developmental pathways and have subtle effects that contribute to health-related phenotypes. These impacts are not well understood.

NHGRI and NICHD recently launched dGTEx to characterize transcriptional profiles in developing humans. dGTEx is a follow-on to the Common Fund Genotype-Tissue Expression (GTEx) Project, which established gene expression profiles within and between human adults across multiple tissues, determining tissue-specific expression as well as correlating differences in expression with genetic variation. dGTEx will study gene expression patterns in multiple human tissues during post-natal and juvenile development, defining the role of gene expression in human development and providing insight into health- and disease-related processes that have their origins before adulthood. While these human studies are valuable, they have some significant limitations, including: 1) data collection in dGTEx is limited to post-natal timepoints; 2) post-mortem studies in humans are necessarily opportunistic in terms of time and cause of death; and 3) human studies involve substantial uncontrolled variability due to environmental and lifestyle differences as well as post-mortem time to tissue sample collection. Use of model organism systems can overcome these constraints and allow for studies at precise developmental stages—both prenatal and postnatal—while controlling for environmental factors.

Evolutionary comparisons have been useful for understanding the effect of genomic variation on biological processes and phenotypes, providing insights about the interplay of genomic variants and environmental pressures and the relevance of putative pathogenic variants identified in clinical studies. As emphasized in the NHGRI Strategic Vision, genomic data from multiple species have been instrumental in elucidating the consequences of natural genomic variation, the conservation of genomic elements, and the rapid evolutionary changes in genomic regions associated with specific traits. Comparison across species at different evolutionary distances can help to identify functional constraints and facilitate correlation of genomic differences with specific traits and phenotypes.

Many model organisms have been used in developmental biology studies, each with different strengths and weaknesses. NHPs have developmental trajectories that are most similar to humans and therefore are used as models to identify genomic contribution to traits and conditions that are highly developed in humans, such as learning and adaptive immunity. Having comparable data from two NHP species at different evolutionary distances from humans, for example, representatives of Old World and New World primates, can enable comparative genomic analyses of development. NHPs are experimentally tractable, providing the opportunity to validate observational findings with follow-up experiments. Unlike many other model systems, they also have high levels of population genomic diversity even in captive populations, making them good models for humans. Therefore, to further explore the impact of genomic variation on development, we propose a comparative genomics approach to developmental tissue expression studies in NHPs to complement the dGTEx initiative in humans.

Proposed Scope and Objectives

The major goal of this proposal is to study gene expression patterns in multiple reference tissues across developmental stages in NHPs and compare them to human gene expression patterns. This will be achieved through two objectives:

Create a tissue resource and generate gene expression data in developing non-human primates by: a) collecting and banking tissues from multiple animals at a minimum of 6 developmental stages spanning prenatal and postnatal development for at least two NHP species used in biomedical research; and b) performing genome sequencing on each animal and RNA sequencing on whole tissues and single cell populations for at least 30 tissue types per animal. These data should be collected with design considerations that maximize their utility to validate and interpret the human data generated in dGTEx (e.g., assaying a combination of comparable and complementary tissues and developmental stages; coordinated sample collection and preservation procedures; standardized quality control/quality assessment, metadata and common data elements) as well as to provide reference data for follow-up functional experiments and experimental interventions in NHPs.

Make the NPH tissue resource and biobank data available and useable to the community by: a) monitoring study progress and laboratory performance; b) performing quality control and basic analysis of expression data; c) identifying and implementing standards for data and metadata; and d) submitting datasets to appropriate resources. In order to maximize impact, it will be important to leverage and integrate the data from this effort with existing developmental genomic data from humans and other resources.

To achieve these objectives, NHP dGTEx will collect and assay samples from at least two different species representing Old World and New World primates for at least 6 developmental stages each with at least 12 individuals at each timepoint. Prenatal stages should include early-, mid-, and late-gestation and postnatal timepoints which should

correspond to developmental stages (neonatal, early childhood, pre- and post-pubertal) being collected in dGTEx. Samples in postnatal periods should include at minimum 30 tissues per donor from tissues comparable to those selected in dGTEx; plans to select samples in prenatal periods should be designed to capture the most relevant tissues for each stage of development. Animals should be selected to maximize genetic diversity within species and provide equal representation of each sex at each timepoint. Genome sequencing will be performed on blood samples from each animal, along with RNA sequencing on a proposed subset of tissue samples. Additional assays such as single cell RNA sequencing, chromatin accessibility, histone modification, and DNA methylation may also be proposed if well justified. Finally, data integration, visualization, accessibility and analyses must be addressed, as these will be critical to making the resource available and providing broad utility. The awardees will be expected to coordinate with the human dGTEx Tissue Procurement Centers managed by NICHD and dGTEx LDACC (Laboratory, Data Analysis, and Coordinating Center) managed by NHGRI. NHP dGTEx will be managed as part of the dGTEx consortium and awardees will be expected to work closely with other dGTEx components. In addition, NHP dGTEx awardees will need to work closely with existing resources (see below) to make the data widely accessible and useful to a variety of communities.

Solicitations for statistical analysis of NHP data, including development of new analysis methods, will be proposed in future years and would be coordinated with the planned solicitation for dGTEx data analysis and analytical methods development. It is expected that many of the proposed analyses, methods and tools would have utility for both human and NHP resources. In addition, development of comparative analyses methods and tools considering human, NHP, and other model organisms will be encouraged, with additional funds going to support those efforts.

Relationship to Ongoing Activities

As described above, the proposed **NHP dGTEx** will leverage and complement **GTEx** and **dGTEx**. It will add developmental data to the NIH **NHP Resources**, which are used for disease studies. NHP dGTEx would also contribute new data standardization, integration, and comparative genomic analysis tools, enhancing the value of NHP resources.

The **Alliance of Genomic Resources (AGR)** integrates cross-species genomic data and resources from the **Model Organism Databases (MODs)**. Several of the MODs have developmental and functional data and annotations, and AGR plans to include other species and additional data types moving forward. NHP dGTEx awardees would be expected to confer with the AGR to integrate NHP and human data with other relevant resources.

Finally, the NHGRI **Genomic Analysis and Visualization and Informatics Labspace (AnVIL)** provides a computation ecosystem for genomics research. NHP dGTEx datasets would be housed in AnVIL along with GTEx and dGTEx data.

Mechanism of Support
The cooperative agreement mechanism will be used to support 1-2 centers to perform NHP tissue procurement, data generation, integration, and analyses.

Funds Anticipated
NHGRI will commit ~$3.5M/yr over 5 years beginning in FY2022 for a total of $17.5M. Co-funding will be sought from other NIH Institutes and Centers, as well as other agencies interested in comparative developmental genomics.