



The Genotype-Tissue Expression (GTEx) Project

Council of Councils

Susan Koester, PhD – NIH/NIMH

June 19, 2015



The Challenge

How do we translate new genomic findings into clinical targets?

Trait-associated
DNA variant

ACGG**G**CAATCA**C**GT
ACGG**G**CAATCA**C**GT
ACGG**G**CAATCA**C**GT

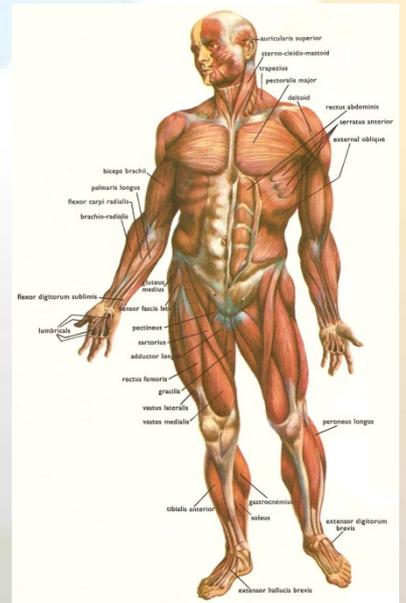
ACGGACAAACAAGT
ACGGACAAACAAGT
ACGGACAAACAAGT



Open questions

→
Causal gene/s?
Causal mechanism?
Causal tissue/s?
Causal pathway/s?

Complex disease



The Opportunity

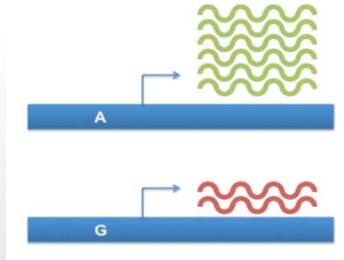
Beyond the GWAS era

Trait-associated
DNA variant

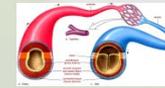
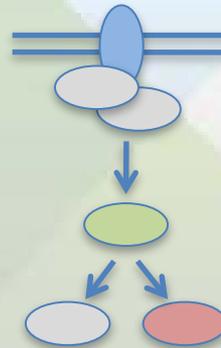
ACGG**G**CAATCAC**C**GT
 ACGG**G**CAATCAC**C**GT
 ACGG**G**CAATCAC**C**GT
 ACGGACAAACAAGT
 ACGGACAAACAAGT
 ACGGACAAACAAGT



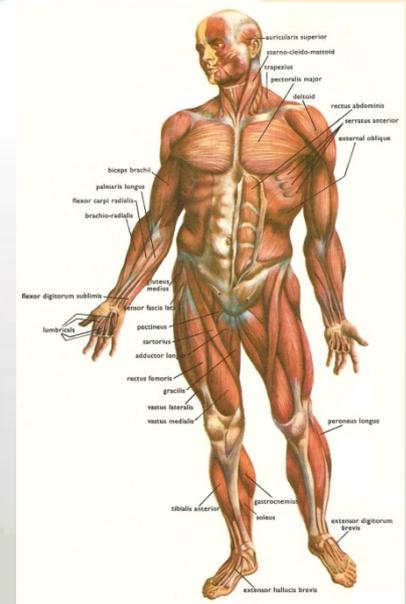
eQTL



Causal genes?
 Causal mechanisms?
 Causal tissues?
 Causal pathways?



Complex disease



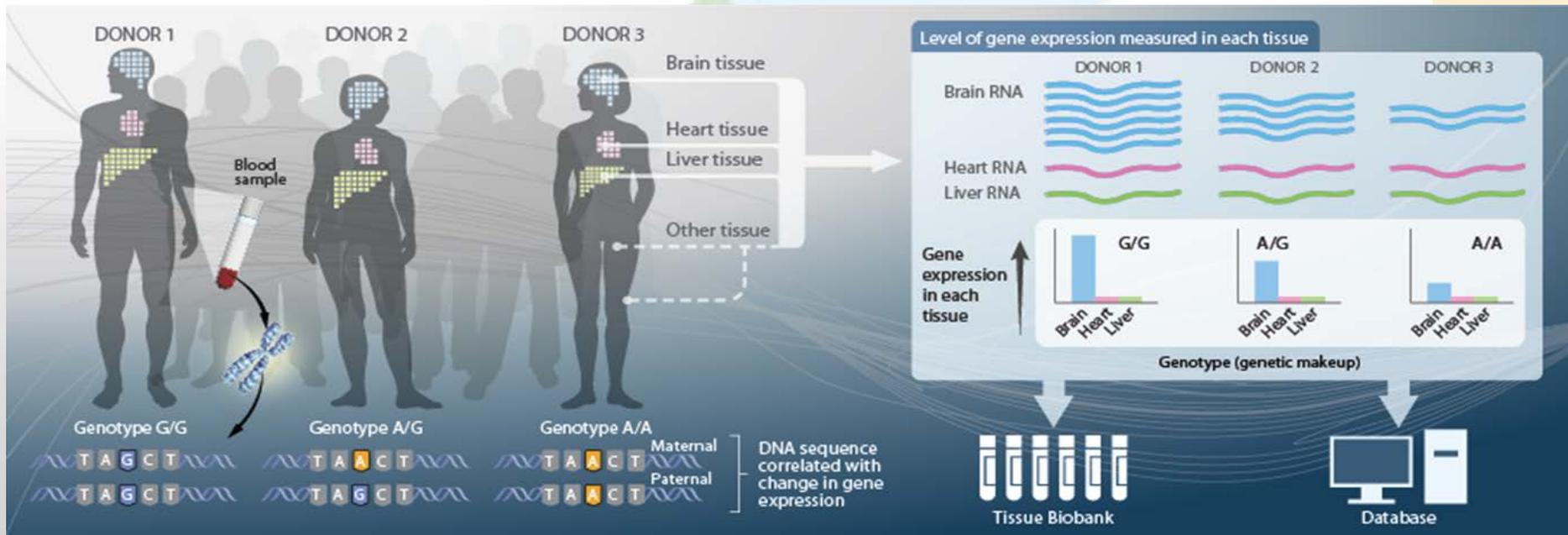
Hypothesis: Disease-associated variants in noncoding regions may be affecting disease through gene regulation

Challenges in using eQTLs to interpret disease associations

- Measuring eQTLs in disease-relevant tissues or cell types
- Most human tissue types are hard to obtain
- Large sample sizes are required for statistical power

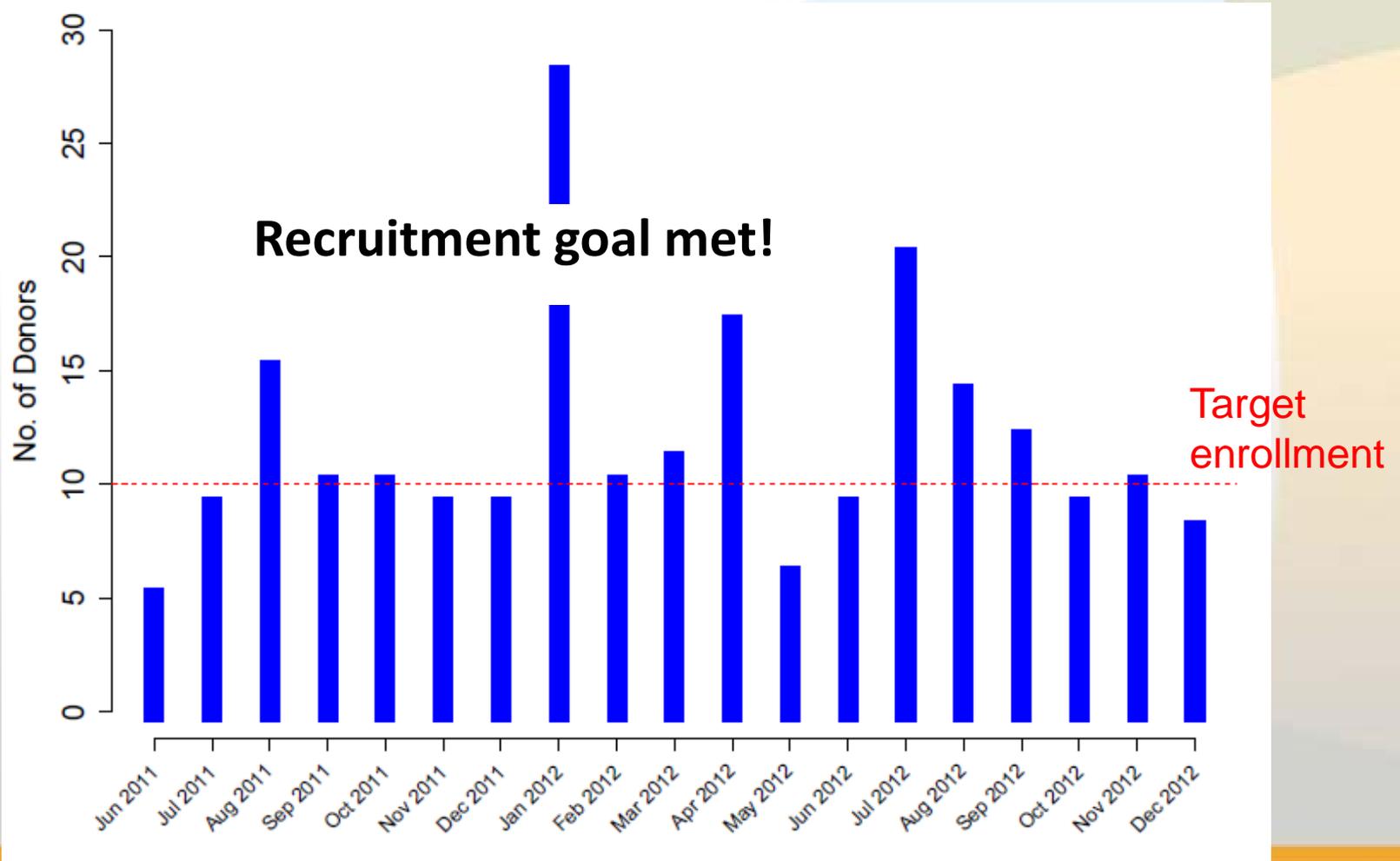
GTEEx Goal

- Help unravel the complex interplay between genetic variation and gene expression across a wide range of non-diseased human tissues.
 - Collect over 30 tissues in rapid autopsy setting
 - WES & WGS and RNA-Seq



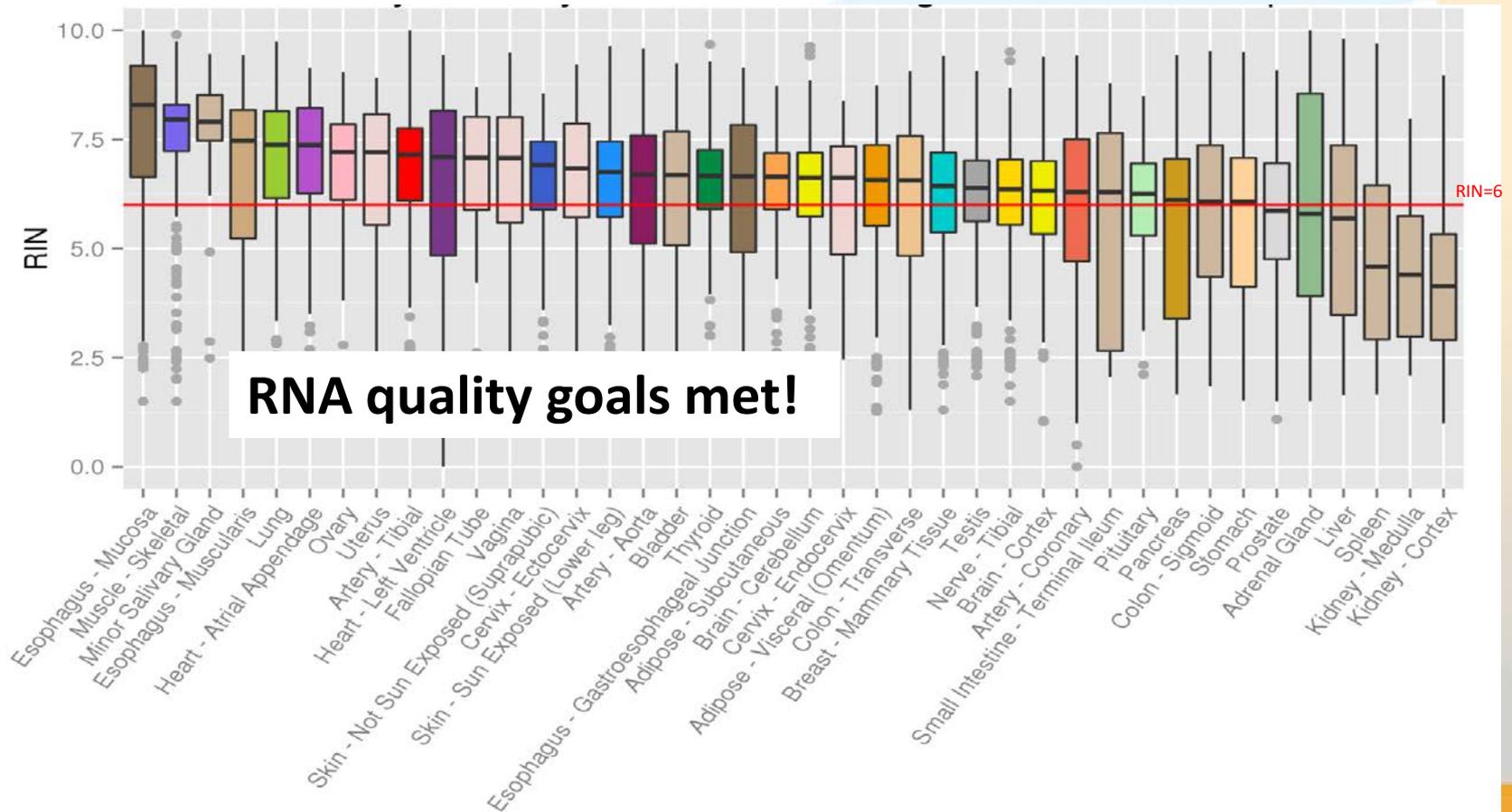
GTEx Pilot

1) Enroll at least 10 post mortem donors per month by the end of the pilot period



GTEx Pilot (cont.)

- 2) Obtain high-quality RNA, defined as a RNA Integrity Number (RIN) of > 6 for 70% of 12 or more organs



GTEx Scale up

- The “Atlas of Human Gene Expression”
- Comprehensive *cis*- and *trans*-eQTL results
 - 900 post-mortem donors, completely sequenced
 - Over 25,000 tissues
 - Gene expression (RNA-Seq) for >20,000 tissue samples
 - Associated clinical and histopathological information
 - Access system for data and samples
 - ELSI study of donor families
 - Beyond Gene Exp

OPO/PM Donors

Goal n=900



From each donor



NIH- supported brain bank

24 hrs

Flash Frozen (Entire brain)



Goal: 30+ tissues
(Total n > 25,000 aliquots)

For each tissue

For subset of tissues



Fix in PAXGene Tissue
(0.2 – 0.5 gram aliquots)

Flash frozen



BSS ↑

CBR

Histopathologic review

1. PAXgene Fixed Paraffin Embedded (PFPE)
2. One aliquot sent out

24 hrs

LDACC ↓

Homogenize, extract DNA & RNA,
measure RNA quality

DNA - Stored

High-quality RNA

fibroblast cell line

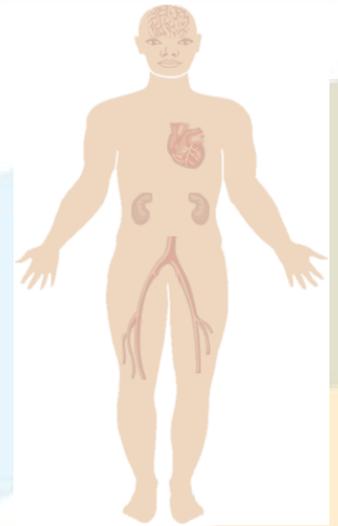
WES / WGS

&
lymphoblastoid cell line

RNA-Seq to 70 million reads; 76 bp paired-end Illumina HiSeq2000

BSS – Biospecimen Source Site
CBR – Comprehensive Biospecimen Resource
LDACC – Laboratory, Data Analysis, and Coordinating Center

Donor Selection Criteria



- Any racial and ethnic group and sex
- Age 21-70
- Collection can start within 24hr of death
 - Organ/Tissue Donors (OPO) & Postmortem Donors (PM)
- Few medical exclusionary criteria:
 - HIV infection or high-risk behaviors, viral hepatitis, metastatic cancer, chemotherapy or radiation therapy for any condition within the past 2 years, whole blood transfusion in past 48 hours, or body mass index ≥ 35 or ≤ 18.5
 - Brain collected if not on ventilator last 24 hours

Recent Publications



RESEARCH ARTICLE

HUMAN GENOMICS

The Genotype-Tissue Expression (GTEx) pilot gene regulation

The GTEx Consortium^{1,†}

Understanding the function of the complex human disease genome is a major goal of biomedical research. We present a catalog of variation in human gene expression across 43 tissues from 1 Genotype-Tissue Expression (GTEx) pilot gene regulation study and identify signals from

These findings provide a systematic understanding of the cellular and biological consequences of human genetic variation and of the heterogeneity of such effects among a diverse set of human tissues.

REPORTS

HUMAN GENOMICS

The human transcriptome in 45 tissues and cell lines

Marta Melé,^{1,2,*} Pedro Monteiro,^{1,7,9} Dmitri D. Pervouchin,^{1,7} Sarah Djebali,^{1,7} An Tuuli Lappalainen,³ Kristin G. Ardlie,^{8,†}

Transcriptional regulation of organismal phenotype by Genotype-Tissue Expression (GTEx) pilot gene regulation study in 45 tissues and cell lines

RESEARCH | REPORTS

HUMAN GENOMICS

Effect of predicted protein-truncating genetic variants on the human transcriptome

Manuel A. Rivas,^{1,*†} Matti Pirinen,^{2,†} Donald F. Conrad,^{3,†} Monkol Lek,^{4,5,†} Emily K. Tsang,^{6,7,8} Konrad J. Karczewski,^{4,5} Julian B. Maller,^{4,5} Kimberly R. Kukurba,^{6,7} David S. DeLuca,⁴ Menachem Fromer,^{4,5,9} Pedro G. Ferreira,^{10,11,12} Kevin S. Smith,^{6,7} Rui Zhang,⁶ Fengmei Zhao,^{4,5} Eric Banks,⁴ Ryan Poplin,⁴ Douglas M. Ruderfer,^{9,13} Shaun M. Purcell,^{4,5,9,13} Taru Tukiainen,^{4,5} Eric V. Minikel,^{4,5} Peter D. Stenson,¹⁴ David N. Cooper,¹⁴ Katharine H. Huang,⁴ Timothy J. Sullivan,⁴ Jared Nedzel,⁴ The GTEx Consortium, The Geuvadis Consortium, Carlos D. Bustamante,⁶ Jin Billy Li,⁶ Mark J. Daly,^{4,5} Roderic Guigo,¹⁵ Peter Donnelly,^{1,16} Kristin Ardlie,⁴ Michael Sammeth,^{15,17} Emmanouil T. Dermitzakis,^{10,11,12} Mark I. McCarthy,^{1,18} Stephen B. Montgomery,^{6,7} Tuuli Lappalainen,^{6,10,11,12,19,20,*†} Daniel G. MacArthur^{4,5,21,*†}

Transcriptional regulation of organismal phenotype by Genotype-Tissue Expression (GTEx) pilot gene regulation study in 45 tissues and cell lines. These signatures are dominated by a relatively small number of genes—which is most clearly seen in blood—though few are exclusive to a particular tissue and vary more across tissues than individuals. Genes exhibiting high interindividual expression variation include disease candidates associated with sex, ethnicity, and age. Primary transcription is the major driver of cellular specificity, with splicing playing mostly a complementary role; except for the brain, which exhibits a more divergent splicing program. Variation in splicing, despite its stochasticity, may play in contrast a comparatively greater role in defining individual phenotypes.

Recent Publications (cont.)



GENOME
RESEARCH

The land



Authors

Yael Bar
Manuel A.
Kukurba
GTEx (C)
Burchard
Montgon

Bioinformatics Advance Access published May 7, 2015

Bioinformatics, 2015, 1–8

doi: 10.1093/bioinformatics/btv074

Advance Access Publication Date: 27 March 2015

Original paper

OXFORD

Gene expression

Assessing allele-specific expression across multiple tissues from RNA-seq read data

Matti Pirinen^{1,*}, Tuuli Lappalainen^{2,3,4,5,6,7}, Noah A. Zaitlen⁸,
GTEx Consortium, Emmanouil T. Dermitzakis^{2,3,4}, Peter Donnelly^{9,10},
Mark I. McCarthy^{9,11} and Manuel A. Rivas^{9,*}



Working Group Members

- **Co-Chairs:**

Eric Green, M.D., Ph.D. Director, NHGRI
Thomas Insel, M.D. Director, NIMH

- **Working Group Coordinators:**

Susan Koester, Ph.D., Deputy Director, DNBB, NIMH
Simona Volpi, Ph.D., Program Director, DGM, NHGRI

- **Project Team Leaders:**

A. Roger Little, Ph.D., Deputy Director, DNBDR, NIDA
Helen M. Moore, Ph.D., Chief, BBRB, NCI
Jeffery P. Struewing, M.D., M.S., Program Director, DGM, NHGRI

- **Common Fund Program Leader:**

Concepcion Nierras, Ph.D., Program Leader, OSC, DPCPSI , OD

- **Members:**

Anjene Addington, Ph.D., Program Officer, NIMH
James F. Battey, M.D., Ph.D., Director, NIDCD
Joy Boyer, Senior Program Analyst, ELSI, NHGRI
Phil Branton, M.D., Pathologist, NCI
Latarsha Carithers, Ph.D., Project Manager, BBRB, NCI
Danielle Carrick, Ph.D., Program Director, NCI
Arthur Castle, Ph.D., Program Director, NIDDK
Christine Colvis, Ph.D., Director, NCATS
Michael Feolo, Ph.D., Staff Scientist, dbGaP, NCBI

Robert Finkelstein, Ph.D. Associate Director, DER, NINDS
Colin Fletcher, Ph.D., Program Director, NHGRI
Weiniu Gan, Ph.D., Program Director, NHLBI
Mark Guyer, Ph.D., Contractor, NHGRI
Emily Harris, Ph.D., M.P.H., Chief, TTGRB, NIDCR
Teresa L.Z. Jones, M.D., Program Director, NIDDK
Alice Kau, Ph.D., HAS, NICHD
Walter J. Koroshetz, M.D., Deputy Director, NINDS
Thomas Lehner, Ph.D., M.P.H., Chief, GRB, NIMH
Nicole Lockhart, Ph.D., Program Director, ELSI, NHGRI
Douglas Meinecke, Ph.D., HSA , NIMH
Alan Michelson, M.D., Ph.D., Associate Director, NHLBI
Stephanie J. Murphy, V.M.D., Ph.D., DAACLAM, Director, DCM ORIP, OD
Melissa Parisi, M.D., Chief, IDDB, NICHD
Jonathan D. Pollock, Ph.D., Chief, GMNRB, NIDA
Abhi Rao, Ph.D., Program Manager, BBRB, NCI
Zhaoxia Ren, Ph.D., Medical Director, NICHD
Laura L. Rodriguez, Ph.D., Director, OPCE, NHGRI
Winifred Rossi, M.A., Deputy Director, DGCG, NIA
John S. Satterlee, Ph.D., Program Director, NIDA
Geetha Senthil, Ph.D., Program Director, NIMH
Nataliya Sharopova, Ph.D., Staff Scientist, dbGaP, NCBI
Pothur Srinivas, Ph.D., HAS, NHLBI
Anne Sturcke, Ph.D., Staff Scientist, dbGaP, NCBI
Anna Taylor, Ph.D., HPS, NINDS
Tiffany Yates, Program Analyst, NIMH
Cuilin Zhang, M.D., M.P.H., Ph.D., Investigator, Epidemiology Branch, NICHD

