

The Genotype-Tissue Expression (GTEx) Project

Latarsha J. Carithers and Helen M. Moore

THIS ISSUE OF *BIOPRESERVATION AND BIOBANKING* highlights the work of the U.S. National Cancer Institute (NCI) and its partners in collecting high quality post-mortem biospecimens for a major National Institutes of Health (NIH) research program, the Genotype-Tissue Expression (GTEx) project.¹ GTEx aims to evaluate the relationship between genetic variation and gene expression in normal human tissues and ultimately understand how this relationship contributes to disease susceptibility and development. GTEx planned to generate gene expression profiles from over 30 different non-diseased (normal) tissue types from hundreds of donors in order to make associations between disease phenotypes and SNPs identified in genome-wide association studies (GWAS). The data would be analyzed by the GTEx Consortium, an international group of investigators funded by the NIH. In the process, all of the data would be available to researchers through the National Center for Biotechnology Information's database of Genotypes and Phenotypes (dbGaP), and all residual biospecimens would be made available for research.

A major challenge was posed by the project's requirements for a diverse set of tissue types from which high-quality RNA could be isolated and characterized. Collections of "normal" or non-diseased tissues can be difficult to find, and when available they may be reserved as control samples for disease studies. For GTEx, the diverse tissue types needed to be collected from individual donors with matched blood for isolation of germline DNA. Where could the NCI obtain tissues in sufficient volumes that were not affected by disease? Rapid autopsy and therapeutic transplant programs run by Organ and Tissue Procurement Organizations (OPOs) and their partner institutions provided a solution for GTEx. Such organizations are expert at working with postmortem donors and their families for transplant purposes as well as for research tissue donation; multiple tissue types could be collected from each donor therefore meeting the demands of the project.

In the article by Carithers, et al., in this issue, the authors describe the robust biospecimen collection infrastructure they developed to support GTEx.¹ Two Biospecimen Sources Sites (BSSs), Roswell Park Cancer Institute and the National Disease Research Interchange (NDRI), managed collections in partnership with OPOs and through support and oversight by NCI and Leidos Biomedical, Inc. To date, biospecimens from almost 900 donors have been collected, including whole brain biospecimens from over 400 donors that are shipped to the University of Miami

Brain Bank for dissection, processing, and storage. Aliquots of all biospecimens are shipped to the Broad Institute of Harvard and MIT for isolation of DNA and RNA and subsequent sequencing, as well as generation of cell lines. Pathology quality control is performed on matching aliquots. The results of the molecular analyses of the first 175 GTEx donors were published in a recent issue of *Science*.² In addition, an embedded study of the ethical, legal, and social implications of GTEx is ongoing³ and Community Advisory Boards have been established to advise best practices in community engagement for postmortem research biospecimen donation.

One of the hallmarks of the GTEx biospecimen collection infrastructure is the emphasis on establishing standard operating procedures (SOPs) for all operational aspects of the project and monitoring adherence to SOPs. This attention to quality control of biospecimens and associated data provides a high level of confidence in the resulting molecular data – an invaluable asset for a project of GTEx' size that aims to produce foundational DNA and RNA sequence data. The GTEx data is already being utilized in multiple studies of human disease and is expected to be useful for many years into the future. Resources generated from GTEx and available to researchers include: the DNA and RNA sequence data available at dbGaP; the Expression Quantitative Trait Loci (eQTL) data available at the GTEx portal (<http://www.gtexportal.org/home/>); the residual biospecimens available to researchers through an access policy (<http://www.gtexportal.org/home/samplesPage>); SOPs publicly available at the website for NCI's Biorepositories and Biospecimen Research Branch (BBRB) (<http://biospecimens.cancer.gov/resources/sops/library.asp>); and the Comprehensive Data Resource (CDR), the IT infrastructure built for the project, available as open source code with adoption enabled through a current NCI Collaboration Opportunity (<http://biospecimens.cancer.gov/newsevents/news/07172015.asp>). An online atlas of the histological images for each biospecimen, complete with pathologist annotations, is in development. To share information about GTEx and raise awareness about the resources available, the NCI hosted a symposium in May 2015, "The GTEx Symposium: All things considered- biospecimens, 'omics data, and ethical issues." Video recordings of the symposium can be found at <http://videocast.nih.gov/Summary.asp?file=19024&bhcp=1> and <http://videocast.nih.gov/Summary.asp?file=19026&bhcp=1>.

Biorepositories and Biospecimen Research Branch, Cancer Diagnosis Program, National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, Maryland.

NCI and its partners are developing a legacy plan to ensure that the GTEx biospecimens will be appropriately managed and available to the research community even after GTEx biospecimen collection concludes. Donor families consented to allow the biospecimens to be used in broad future research and therefore the biospecimens will be useful for years to come. We thank the families of the donors for their generous gift and acknowledge their contribution to medical research.

Postmortem biospecimen collection can result in invaluable samples for research. This issue of the *Journal* also highlights the 35th anniversary of the National Disease Research Institute (NDRI) in an article entitled, “Times are changing: 35 years of human biospecimen procurements for the National Disease Research Interchange.”⁴ We congratulate NDRI on this important milestone and acknowledge the contribution of their biospecimen collection programs to medical research, including their critical role in GTEx. The scientific requirements for GTEx presented complex and significant challenges for biospecimen collection. With outstanding efforts put in by NIH, its major partners, and the many members of the GTEx Consortium, those challenges have been met and a high quality biospecimen collection created. The strength and utility of the resulting molecular data has yet to be fully realized, as sequencing and analysis is ongoing and expected to be completed in 2017. We are confident that the results will be worth the incredible effort and we thank all involved for their dedication to this groundbreaking project.

1. Carithers LG, Ardlie K, Branton PA et al. on behalf of the GTEx Consortium: A novel approach to high-quality post-mortem tissue procurement: The GTEx Project. *Biopreserv Biobank* 2015; 13:307–308.
2. The GTEx Consortium: Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* 2015; 348:648–660.
3. Siminoff LA, Traino HM, Mosavel M, Barker L, Gudger G, Undale A: Family decision maker perspectives on the return of genetic results in biobanking research. *Genet Med* 2015. Epub ahead of print 9 April 2015.
4. Bell TG, Leinweber B: Times are changing: 35 years of human biospecimen procurements for the National Disease Research Interchange. *Biopreserv Biobank* 2015; 13:309–310.

Address correspondence to:

Helen M. Moore, Ph.D.

*Chief, Biorepositories and Biospecimen Research Branch
Cancer Diagnosis Program*

*Division of Cancer Treatment and Diagnosis
National Cancer Institute, Rm 3W428, MSC 9728
9609 Medical Center Drive
Bethesda, MD 20892*

E-mail: moorehe@mail.nih.gov