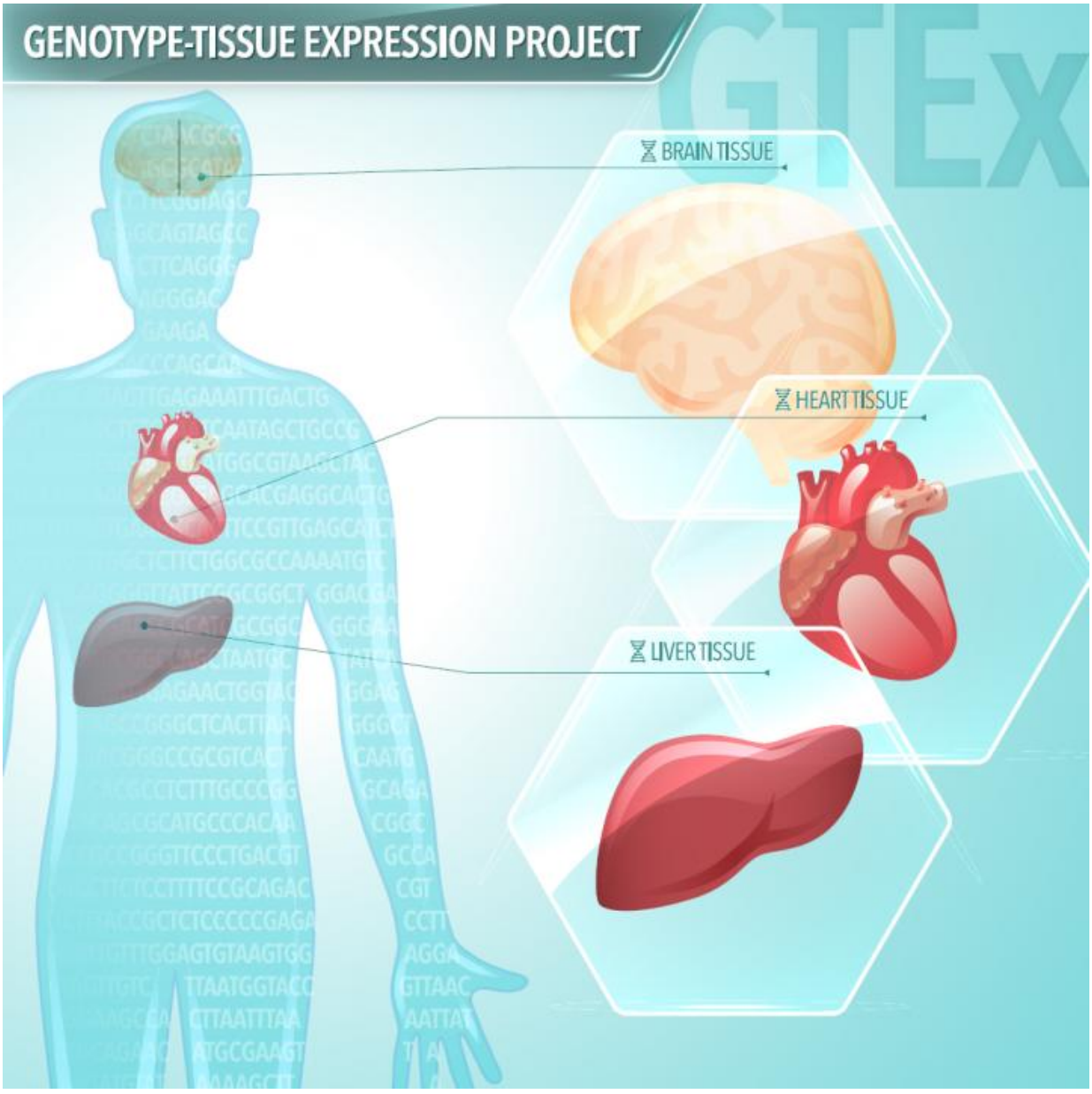
New GTEx findings show how DNA differences influence gene activity, disease susceptibility

May 7 2015



GTEx findings reveal new insights into how DNA differences influence gene activity, disease susceptibility. Credit: Ernesto Del Aguila, NHGRI

Researchers funded by the National Institutes of Health Genotype-Tissue Expression (GTEx) project have created a new and much-anticipated data resource to help establish how differences in an individual's genomic make-up can affect gene activity and contribute to disease. The new resource will enable scientists to examine the underlying genomics of many different human tissues and cells at the same time, and promises to open new avenues to the study and understanding of human biology.

GTEx investigators reported initial findings from a two-year pilot study in several papers appearing online May 7, 2015, in *Science* and other journals. These efforts provide new insights into how genomic variants - inherited spelling differences in the DNA code - control how, when and how much genes are turned on and off in different tissues, and can predispose people to diseases such as cancer, heart disease and diabetes.

"GTEx was designed to sample as many tissues as possible from a large number of individuals in order to understand the causal effects of genes and variants, and which tissues contribute to predisposition to disease," said Emmanouil Dermitzakis, Ph.D., professor of genetics at the University of Geneva Faculty of Medicine, Switzerland, and a corresponding author on the main *Science* paper. "The number of tissues examined in GTEx provides an unprecedented depth of genomic variation. It gives us unique insights into how people differ in gene expression in tissues and organs."

NIH launched the GTEx Project in 2010 to create a data resource and tissue bank for scientists to study how genomic variants may affect gene activity and disease susceptibility. Investigators are collecting more than 30 tissue types from autopsy and organ donations in addition to tissue transplant programs. The DNA and RNA from those samples are then analyzed using cutting-edge genomic methods. The project will eventually include tissue samples from about 900 deceased donors. GTEx is supported by the NIH Common Fund and administered by the National Human Genome Research Institute (NHGRI), the National Institute of Mental Health (NIMH) and the National Cancer Institute (NCI), all part of NIH.

"GTEx will be a great resource for understanding human biological function, and will have many practical applications in areas such as drug development," said NHGRI Program Director Simona Volpi, Pharm.D., Ph.D. "Scientists studying asthma or kidney cancer, for example, will be interested in understanding how specific variants influence the biological function of the lung, kidney and other organs."

"Projects supported by the Common Fund aim to advance multiple areas of biomedical research," said James M. Anderson, M.D., Ph.D., director of the NIH Division of Program Coordination, Planning, and Strategic Initiatives, which houses the Common Fund. "The unprecedented breadth of GTEx donors and tissue types establishes a resource that scientists studying areas ranging from blood pressure to neurodegenerative disease would find invaluable."

In the main *Science* paper, researchers analyzed the gene activity readouts of more than 1,600 tissue samples collected from 175 individuals and 43 different tissues types. One way that researchers evaluate gene activity is to measure RNA, which is the readout from the genome's DNA instructions. Investigators focused much of their analyses on samples from the nine most available tissue types: fat, heart,

lung, skeletal muscle, skin, thyroid, blood, and tibial artery and nerve.

The genomic blueprint of every cell is the same, but what makes a kidney cell different from a liver cell is the set of genes that are turned on (expressed) and off over time and the level at which those genes are expressed. GTEx investigators used a methodology - expression quantitative trait locus (eQTL) analysis - to gauge how variants affect gene expression activity. An eQTL is an association between a variant at a specific genomic location and the level of activity of a gene in a particular tissue. One of the goals of GTEx is to identify eQTLs for all genes and assess whether or not their effects are shared among multiple tissues.

Investigators discovered a set of variants with common activity among the different tissue types. In fact, about half of the eQTLs for protein-coding genes were active in all nine tissues. They identified approximately 900 to 2,200 eQTL genes - genes linked to nearby genomic variants - for each of the nine tissues studied, and 6,486 eQTL genes across all the tissues. "We didn't know how specific this regulation would be in different tissues," said co-corresponding author Kristin Ardlie, Ph.D., who directs the GTEx Laboratory Data Analysis and Coordination Center at the Broad Institute of MIT and Harvard in Cambridge, Massachusetts. "The analysis showed a large number of variants whose effects are common across tissues, and at the same time, there are subsets of variants whose effects are tissue-specific."

Comparing tissue-specific eQTLs with genetic disease associations might help provide insights into which tissues are the most relevant to a disease. The researchers also found a great deal of eQTL sharing among tissues, which can help explain how genomic variants affect the different tissues in which they are active.

Even when active in multiple tissues, the same variant can sometimes

have a different effect in different tissues. GTEx researchers found, for example, that a variant that affects the activity of two genes associated with blood pressure had a stronger effect on gene expression relevant to blood pressure in the tibial artery - even though there was greater overall gene activity in other tissues. They also noted that the same gene activity profiles characterizing tissues from living donors were seen in the GTEx samples from deceased donors.

Two companion studies in *Science* used GTEx data to examine other aspects of gene activity in different tissues. One study characterized the effects of protein-truncating variants (PTVs) on gene activity. PTVs shorten the protein-coding sequence of genes, and affect their function. Some rare PTVs can lead to diseases, such as Duchenne muscular dystrophy. Each person's genome carries about 100 PTVs, though most have little or no effect (and in some cases can even protect against disease).

Manuel Rivas, a Ph.D. candidate at the University of Oxford, and his colleagues used GTEx data and information from a large European project to examine the gene readouts from more than 600 individuals. The team found PTVs that affect protein production either through the degradation of gene transcripts or by disrupting a process called splicing. In both cases, the researchers were able to use the GTEx data to measure these effects across individuals and tissue types. The group is now developing better methods for predicting the impact of PTVs identified in patients with diseases.

In another companion study in *Science*, Roderic Guigo, Ph.D., coordinator for the Bioinformatics and Genomics Program at the Centre for Genomic Regulation in Barcelona, Spain, and his colleagues examined patterns in gene readouts across nearly 1,500 GTEx tissue samples. The researchers found that gene activity differed substantially more across tissues than across individuals.

Investigators discovered just under 2,000 genes that vary with age, including genes related to neurodegenerative diseases such as Parkinson's disease and Alzheimer's disease. They also found more than 750 genes with differences in activity between men and women. Some genes are related to diseases with differences in prevalence between men and women, including five related to heart disease.

Three other studies analyzing GTEx data also appear May 8 in the journals *Bioinformatics*, *PLoS Computational Biology* and *Genome Research*.

More information: "The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans," www.sciencemag.org/lookup/doi/ … 1126/science.1262110

"The human transcriptome across tissues and individuals," www.sciencemag.org/lookup/doi/ … 1126/science.aaa0355

"Effect of predicted protein-truncating genetic variants on the human transcriptome," www.sciencemag.org/lookup/doi/ … 1126/science.1261877

Provided by NIH/National Human Genome Research Institute