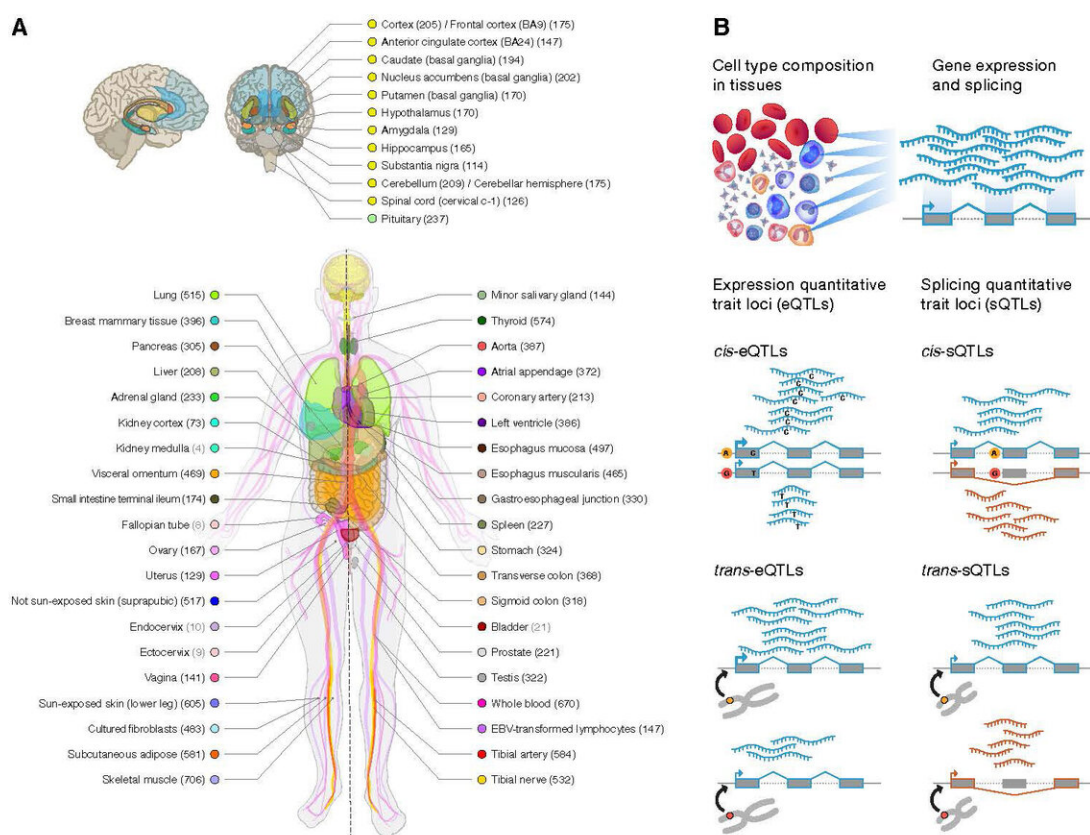


GTEx Consortium releases fresh insights into how DNA differences govern gene expression

September 10 2020



(A) Illustration of the 54 tissue types examined (including 11 distinct brain regions and two cell lines), with sample numbers from genotyped donors in parentheses and color coding indicated in the adjacent circles. Tissues with 70 or more samples were included in QTL analyses. (B) Illustration of the core data types used throughout the study. Gene expression and splicing were quantified

from bulk RNA-seq of heterogeneous tissue samples, and local and distal genetic effects (cis-QTLs and trans-QTLs, respectively) were quantified across individuals for each tissue. Credit: Aguet et al., *Science* (2020)

Scientists from the Genotype-Tissue Expression (GTEx) project, a National Institutes of Health-funded consortium including researchers from the Broad Institute of MIT and Harvard, have completed a wide-ranging set of studies documenting how small changes in DNA sequence can impact gene expression across more than four dozen tissues in the human body.

These studies, released in a set of 15 papers published in *Science* and other journals, constitute the most comprehensive catalog to date of genetic variations that affect [gene expression](#). They also highlight the importance of cell type as a factor in understanding how genes are regulated in [human tissues](#), and provide a rich resource for connecting the functional dots between [genetic variation](#) and human traits and diseases.

The NIH launched GTEx in 2010 to identify and map quantitative trait loci (QTLs), namely, associations between genetic variants at specific locations in the genome and gene expression within a variety of tissues. Researchers have mapped the vast majority of genetic variants discovered through [genome-wide association studies](#)—which scan the genome to identify variants linked to traits or disease—to regions of the genome's non-coding DNA (which does not directly instruct the construction of proteins). This suggests that these variants act by influencing genes' expression, rather than by altering the proteins they encode.

To shed light on these relationships, GTEx set out to genotype and

measure gene expression in samples of up to 50 [tissue types](#) (brain, heart, lung, prostate, uterus, etc.) from as many as 1,000 deceased donors, with a goals of identifying QTLs for as many genes as possible, and determining whether or not their effects are shared among multiple tissues or [cell types](#).

"GTEx attempted to map, across as many individuals as possible, the basis of gene regulation, starting from how a genetic change might affect how a gene is expressed or how a protein is produced," said Kristin Ardlie, who directs the GTEx Laboratory Data Analysis and Coordination Center at Broad, and who served as co-corresponding author on the project's flagship *Science* paper with Broad computational biologist François Aguet and Tuuli Lappalainen of the New York Genome Center (NYGC).

A resource for the future

The flagship *Science* paper details the results of the GTEx Consortium's 10 years of work, efforts that have helped reveal much about the immense complexity underlying genetic control of gene expression. It presents the results of the consortium's analysis of 15,201 samples representing 52 tissues, collected from 838 donors—a dataset nearly twice the size of that behind the most recent prior GTEx papers published in 2017. Each donor underwent whole genome sequencing to identify the genetic variants present, along with RNA sequencing of all [tissue](#) samples to establish the pattern of gene expression within the tissue.

The resulting dataset—available via the GTEx portal—catalogs QTLs governing the expression of more than 23,000 genes, with multiple QTLs regulating many genes. These included variants that directly affect expression of (eQTLs) or splicing within genes (sQTLs), both for variants close to the genes they control (cis-QTLs) and ones located on

chromosomes other than the one harboring their target gene (trans-QTLs).

The data also confirmed that QTLs tend to be either very tissue-specific in their expression effects, or shared quite broadly across all tissues; and revealed some differences in QTL effects between sexes and across populations.

Mechanistically, the findings suggest that QTLs may often affect how a cell's transcription factors bind to the genome at a gene's promoter or enhancer, which in turn affects that gene's expression. And they also provide a baseline for deeper insights into functional roles QTLs play.

"At this larger sample size, and with the diverse tissues and donors we have, we can start to see that there is more than one regulatory effect per gene, and that these differ not just by tissue but by cell type," Ardlie said. "We can start to map at high resolution the variants that actually impact a trait. And we can begin to relate GWAS signals to QTLs and see whether what appear to be random GWAS hits might actually fall within functional elements that affect gene regulation and complex trait and disease phenotypes."

Tuning in

A key focus for this latest set of GTEx studies was to understand how QTLs mapped not just to tissues, but to specific cell types. With hundreds of samples sequenced from many tissues, GTEx researchers found that many [genes](#) were influenced by multiple QTLs. This phenomenon, called "allelic heterogeneity," reflects the fact that the GTEx tissue samples represent mixtures of many types of [cells](#).

To gain a more nuanced understanding of QTLs' cellular specificity and learn the extent to which QTLs from different cell types contributed to

their tissue-level observations, a GTEx team led by Aguet at Broad and Lappalainen and Sarah Kim-Hellmuth at NYGC used the project's RNA profiling data to computationally identify the cell types present within GTEx's tissue samples. They then checked whether QTLs mapped within those tissues were likely to be specific to the inferred cell types.

These analyses, reported in a companion *Science* paper, pinpointed thousands of "cell type interaction QTLs," many of which had not been previously characterized. The results indicate that many more cell type specific QTLs are likely to exist but cannot yet be detected without additional samples or improved methods. They also showed that the patterns of QTL sharing and specificity across tissues could be tied back to whether those tissues shared cell types in common.

The findings also revealed that even at the cell type-level, multiple QTLs can influence any given gene, sometimes acting together to boost expression, sometimes in opposition to tamp expression down, depending on an individual's genotype.

"In a sense, QTLs act like a dial on expression, one that can be adjusted up or down," Aguet explained. "One QTL might increase expression, but another might turn it back down a little. It all adds to the complexity of how genetic variation regulates gene expression."

An end, but also a beginning

This collection of studies comprises the consortium's final analysis of the GTEx dataset, though a great deal of work remains to be done and a great deal of knowledge remains to be gleaned from the catalog of QTLs. For instance, Ardlie noted, QTL analysis provides only one lens through which to view the functional implications of genetic variation, one that complements epigenomic, proteomic, and other forms of genomic and transcriptomic analysis.

"GTEx was an ambitious, complex undertaking, and it remains very difficult to access this breadth of tissues from individuals, and in that sense GTEx was unique and has helped pave the way for studies like the Human Cell Atlas," she said. "But we really need large-scale resources like this and others, such as ENCODE, from which we can glean complementary information to get a more complete picture of the molecular mechanisms that drive biology."

More information: "The GTEx Consortium atlas of genetic regulatory effects across human tissues," by F. Aguet at The GTEx Consortium. *Science* (2020). [science.sciencemag.org/cgi/doi ... 1126/science.aaz1776](https://science.sciencemag.org/cgi/doi/10.1126/science.aaz1776)

Provided by Broad Institute of MIT and Harvard

Citation: GTEx Consortium releases fresh insights into how DNA differences govern gene expression (2020, September 10) retrieved 21 May 2024 from <https://phys.org/news/2020-09-gtex-consortium-fresh-insights-dna.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.
