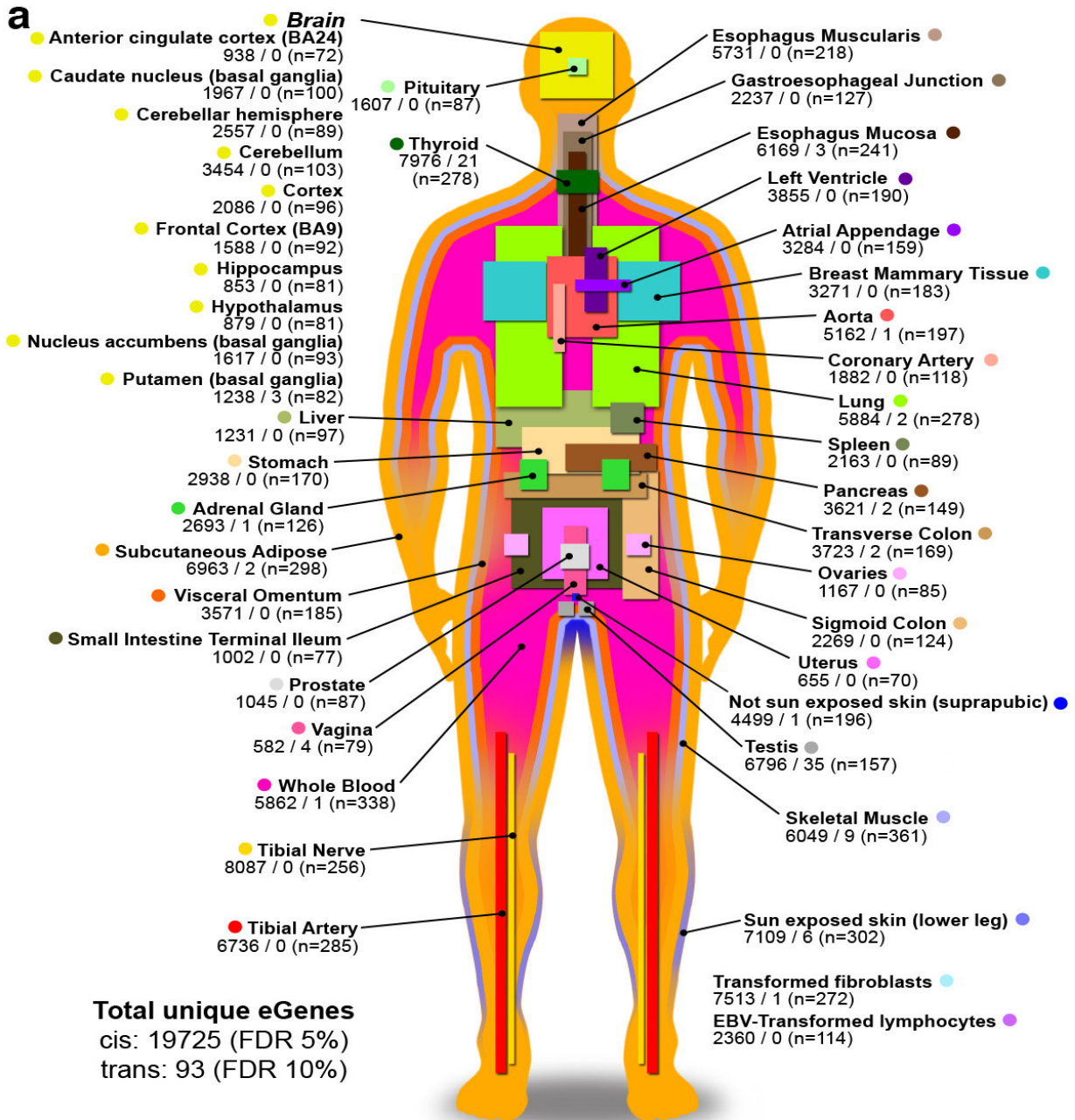


# Scientists demonstrate path to linking the genome to healthy tissues and disease

October 13 2017

---



An international team of researchers reached a major milestone in decoding the human genome by linking genes across all chromosome of many individual people to specific tissues and disease processes. Using tissue samples donated from 449 people, the researchers linked nearly 20,000 genes to 44 tissue types. In the illustration, each tissue type is followed by the number of genes whose level of activity is controlled by nearby genes on the same chromosome (cis);

those whose activity is associated with genes on other chromosomes (trans); and the number of tissue samples studied. Credit: GTEx consortium

Our genomes help to determine who we are - the countless variations between individuals that encode the complexity of tissues and functions throughout the body. Since scientists first decoded a draft of the human genome more than 15 years ago, many questions have lingered, two of which have been addressed in a major new study co-led by a Princeton University computer scientist: Is it possible, despite the complexity of billions of bits of genetic information and their variations between people, to develop a mechanistic model for how healthy bodies function? Furthermore, can this model be used to understand how certain diseases emerge?

On Oct. 11, scientists came the closest yet to delivering an answer of "Yes." An international consortium of researchers in the Genotype-Tissue Expression (GTEx) Consortium published findings about how [genetic variation](#) effects gene regulation in 44 human tissue types. Reported in the journal *Nature*, the data help to establish a baseline understanding of the diversity of genetic roles in maintaining human tissues. The researchers said the work demonstrates that, in fact, multi-tissue, multi-individual data can be used to identify the mechanisms of gene regulation and help to study the genetic basis of complex diseases.

The research that led to these findings is part of a larger effort to better understand gene regulation and [expression](#), carried out by the GTEx Consortium, a National Institutes of Health-funded group that includes researchers from around 80 institutions founded in 2010.

"The ultimate goal is to understand [gene expression](#) and gene regulation in a diversity of tissue types," said Barbara Engelhardt, an assistant

professor in the Department of Computer Science at Princeton University, who is one of four corresponding authors of the paper and a GTEx principal investigator. "This is absolutely critical to understanding how dysregulation may lead to disease."

Scientists are only beginning to reveal, for example, how genetic variation in our 22,000 genes - as well as "non-coding" regions in the genome - help to shape complex traits, from a person's height to whether he or she develops autism. Further, scientists seek to understand interactions between multiple genes and the environment. The same unknowns hold true for how genetic variation contributes to disorders such as schizophrenia and Parkinson's disease.

Teasing apart these complexities first requires characterizing how healthy tissues function, which in turn requires [tissue samples](#). To obtain those samples, GTEx researchers requested consent from family members to collect small pieces of up to 50 different tissues immediately after a donor's death. Samples range from various organs and blood, and include ten brain sub-regions. This work represents data across 449 donors.

"These types of tissue are incredibly difficult to get from healthy living donors," Engelhardt said. "With endless thanks to the donors, we have these samples as a resource. We can now explain observed relationships between genotype and disease by looking at the effects of the genotypes that lead to higher risk of the disease on [gene expression levels](#) in disease-specific tissues, including brain."

While the research is still ongoing, this latest study represents the largest analysis to date, including over 7,000 tissue samples. Engelhardt's group was responsible for mapping associations between genetic variants and gene expression levels on different chromosomes, a connection known as "trans-expression quantitative trait loci (trans-eQTLs)." In contrast,

cis-eQTLs—which account for the majority of genetic variation that affects gene expression—regulate genes located nearby on the same chromosome. Trans-eQTLs in particular have proven especially difficult to identify because of their biological and statistical complexity, Engelhardt said, but they might hold clues for explaining complex traits in a more comprehensive way than cis-eQTLs.

Engelhardt and her group's role in the study included mapping and interpreting trans-eQTLs that they identified in the [tissue](#) samples. After clearing the samples of variance due to technical artifacts that could potentially confound the findings, they performed 3.5 trillion statistical tests against every mutation in the genome compared to every expressed gene in each of the 44 tissues. They used additional statistical techniques to correct for false positives in the data, which left them with several hundred trans-eQTLs. In the study, they additionally confirmed that nearby genetic variation in the form of cis-eQTLs affected expression of about 50 percent of genes in the samples. This work suggests, however, that this figure will climb to closer to 100 percent when more samples are added in the future.

"The extensive catalogue generated by the GTEx consortium takes us one step closer to decoding the regulatory code of the genome," said Yoav Gilad, a geneticist at the University of Chicago who was not involved in the study but was a scientific reviewer on the paper. "The consequences of genetic variation on gene expression are gradually becoming clearer."

One trans-eQTL variant revealed in the study that was of particular interest was a mutation known to increase the risk of thyroid cancer. It is situated just next to a thyroid-specific transcription factor, a protein that regulates the rate of gene expression in the thyroid. Prior to the study, the broad effects of the thyroid-specific transcription factor, called FOXE1, on transcription levels of genes were not well characterized.

The researchers were able to replicate this finding by comparing the healthy thyroid tissues in GTEx to 500 samples taken from thyroid tumors, compiled by The Cancer Genome Atlas, and giving support to the extensive impact of FOXE1 on cellular state.

With these findings, "we can start to think about how to target specific [genes](#) for creating therapies for thyroid cancer," Engelhardt said. "Many thyroid diseases will be impacted by changing the expression levels of the thyroid-specific transcription factor, so we want to investigate FOXE1 more carefully in future work."

While the study represents a strong start for understanding how eQTLs affect [gene regulation](#) and expression, Engelhardt pointed out that she and her colleagues still do not have enough samples to understand trans-eQTLs as deeply as they would like. The GTEx consortium is currently working on an analysis that includes almost three times as many samples as this current study. In addition, they hope to soon extend the project to new, underrepresented populations and build on existing efforts.

"The value of this dataset is in understanding and interpreting results in genome-wide studies," Engelhardt said. "It's already been extremely effective in understanding inherited diseases, and hopefully, as a resource, it continues to improve with more samples an

**More information:** François Aguet et al. Genetic effects on gene expression across human tissues, *Nature* (2017). [DOI: 10.1038/nature24277](https://doi.org/10.1038/nature24277)

Provided by Princeton University

Citation: Scientists demonstrate path to linking the genome to healthy tissues and disease (2017,

October 13) retrieved 26 April 2024 from <https://medicalxpress.com/news/2017-10-scientists-path-linking-genome-healthy.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.