

The GTEx project has created a reference resource of gene expression levels from 'normal', non-diseased tissues. DNA was genotyped at 2.2 million sites and imputed to 12.5 million sites (11.5 million autosomal and 1 million X chromosome sites) using the multi-ethnic reference panel from 1000 Genomes Project Phase 1 v311. For this data release, 44 sampled regions or cell lines were considered, each from at least 70 donors, and thereby considered suitable for eQTL analysis: 31 solid-organ tissues, 10 brain subregions including duplicates of two regions (cortex and cerebellum), whole blood, and two cell lines derived from donor blood and skin samples. All data are available from dbGaP (accession phs000424.v6.p1) with multiple data views publicly available from the GTEx Portal (www.gtexportal.org). Whole-genome sequencing was performed for 148 donors to a mean coverage greater than 30x, and all donors were exome-sequenced to a mean coverage over captured exons of 80x. The resulting data provide the deepest survey of individual- and tissue-specific gene expression to date, enabling a comprehensive view of the impact of genetic variation on gene expression levels. RNA sequencing (RNA-seq) samples were sequenced to a median depth of 78 million reads. RNA was isolated from postmortem samples in an ongoing manner as donors were enrolled into the study. We hereafter refer to these tissues, regions, and cell lines as the 'tissues' used in eQTL analysis. Every tissue sample was examined histologically; the sample was accepted for the project if the tissue was non-diseased and in the normal range for the age of the donor. A total of 7,051 samples from 449 donors represent the GTEx v6p analysis freeze (Fig. Sampled donors were 83.7% European American and 15.1% African American. 1a; Supplementary Information 1-5; Supplementary Figs 1-6; Supplementary Tables 1-10)..