# GTEx Data in the UCSC Genome Browser 2016 Update

Christopher Lee, Kate Rosenbloom, Galt Barber, Parisa Nejad, Luvina Guruvadoo, Cath Tyner, Max Haeussler, John Vivian, Melissa Cline, Benedict Paten, Donna Karolchik, Ann S. Zweig, W. James Kent

*UCSC Genomics Institute*

## Abstract

We have integrated GTEx V6 gene expression data into the UCSC Genome Browser (http://genome.ucsc.edu) and tools, and developed a new display to show tissue-specific expression in genomic context. The new GTEx Gene Expression track was released on the human hg19/GRCh37 and hg38/GRCh38 genome assemblies in April of this year. We have also released a public track hub of read coverage signal tracks for the GTEx V6 samples and incorporated a GTEx expression boxplot into our main gene track details pages. Incorporating summary SNP/Allele-Specific Expression (ASE) level data from GTEx is underway.

## GTEx Gene Expression track display

Genome Browser multi-region display of GTEx gene-level expression data from the V6 analysis shows 3 genes in the leptin signaling pathway. LEP expression induces transcription of SOCS3 and FOS, which in turn induces release of anorexigenic peptides, satiating appetite. For comparison, the ENCODE RNA-Seq from 9 cell types track is also included. Also shown is the **GTEx RNA-Seq Signal Hub**, which displays read coverage for all GTEx V6 samples, organized in tracks by age, sex and tissue.
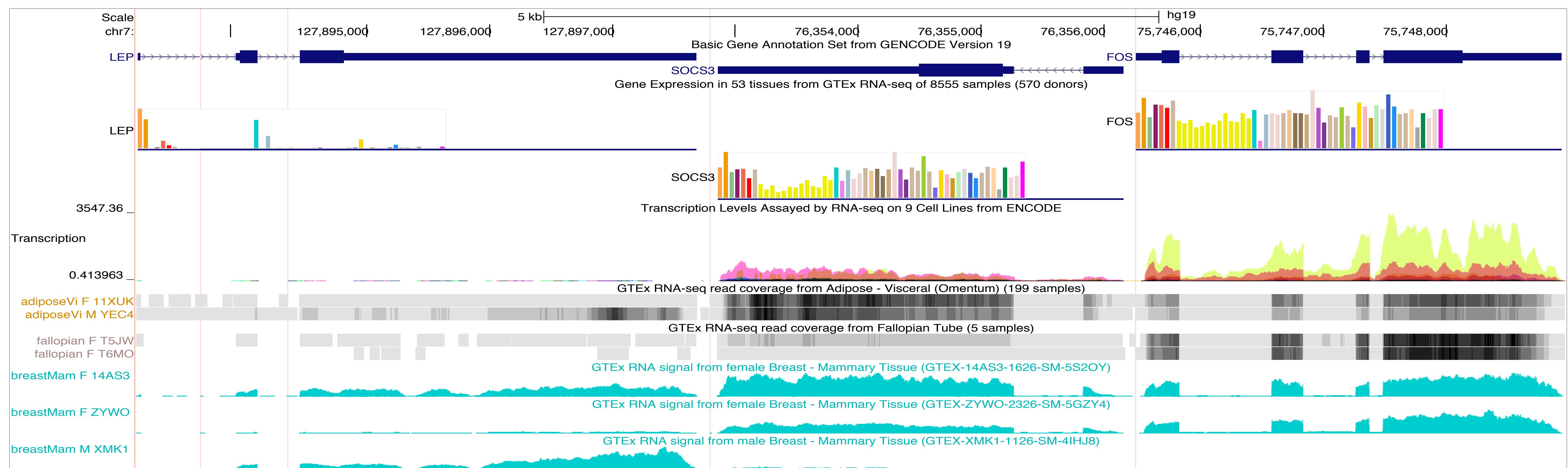


Tissues are colored according to GTEx publication conventions. The display is based on median expression levels in RPKM for each tissue per gene. The displayed expression graphs are fixed-width and are anchored at the genomic start of the GTEx gene model. Thus the orange and turquoise bars in the LEP graph indicate that the entire gene is highly expressed in adipose and breast tissue (rather than indicating high expression of the second exon in breast tissue).

## Track Configuration and Details



- Filter track display via tissue, score, RPKM, or protein coding genes.
- Choose between a linear or $\log_{10}$ transform, similar to the display options from the Broad Institute's GTEx Portal.
- Click through from the main Genome Browser track display to view a detailed boxplot of RPKM and sample number (shown below for the STARD3 gene).
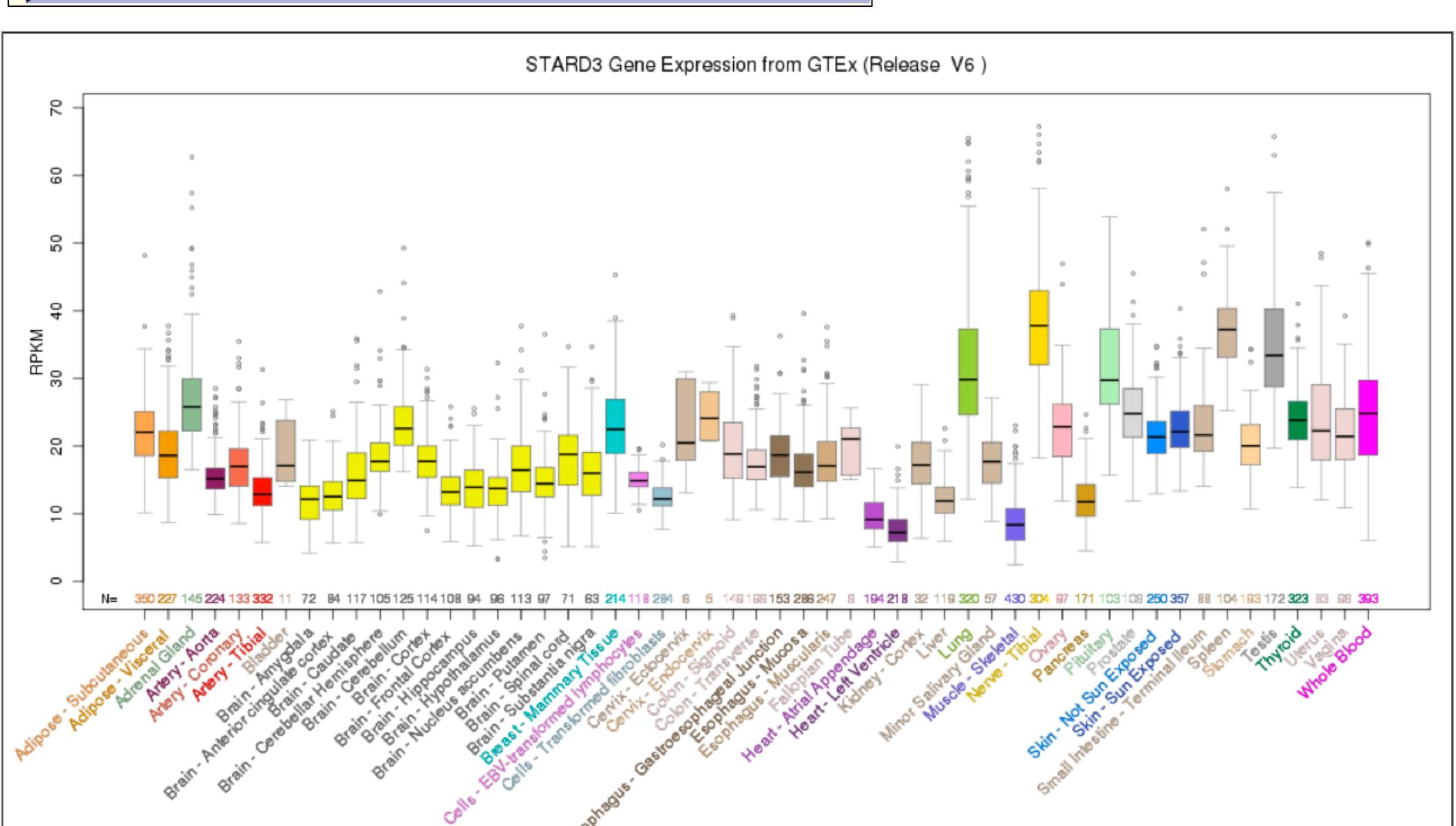


Levels of RNA observed within a gene are available on the hub. Search for "GTEx" on the Track Hubs page (right), to connect to the hub and see the signal graphs. Once connected, explore RNA-Seq data by sex or gender, even configuring which individual samples you would like to display. In the signal tracks shown above, you can see differences in expression of the same tissue in different donors.



Don't forget to zoom out while looking at the GTEx track, as in the ~100kb example below. Here the region upstream of the TP53 gene is displayed, illustrating how the GTEx gene bargraph size is customized for different zoom levels. You can even compare gene expression across large spans of sequence like individual chromosomes!
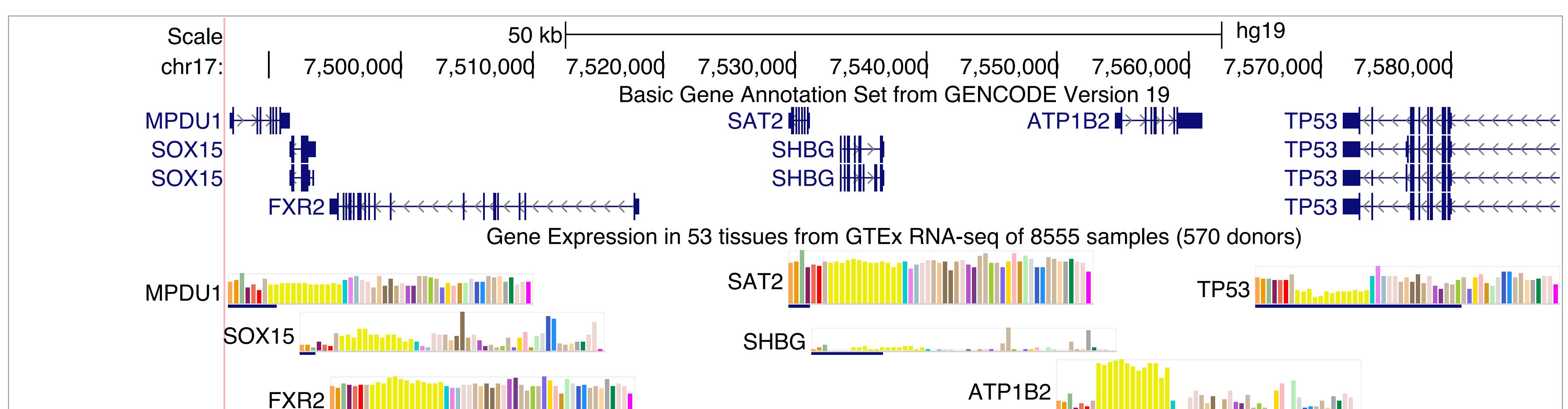


## Mining GTEx Data in the Browser Database

The Table Browser and Data Integrator support queries on GTEx data. On the left below is a Table Browser query for SOCS3 expression in fallopian tube. The query was configured to output tissue, RPKM, gene details, age, and gender. Output from the query is shown on the right.



## Next steps

- Allele specific expression tracks
- Exon expression tracks
- eQTL tracks
- Body Map for tissue selection
- Add GTEx to the Gene Sorter
- Expand data summaries in GTEx track information (e.g. cluster plots).
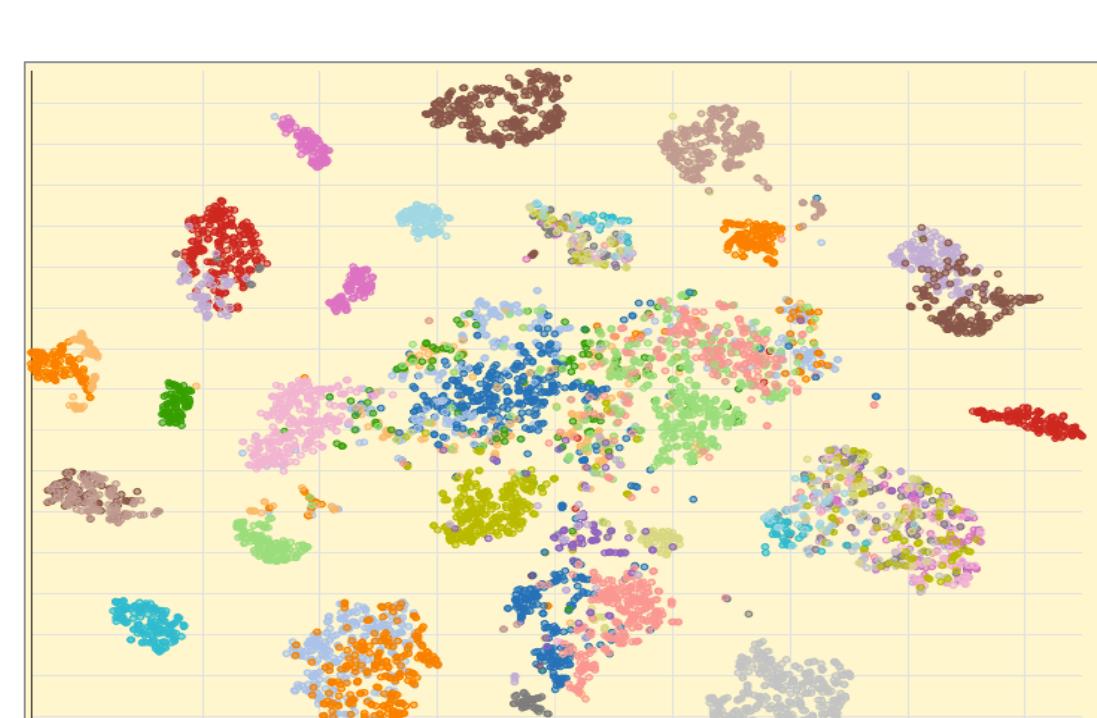


T-SNE clustering of GTEx gene expression, colored by tissue

## Acknowledgements

http://genomewiki.cse.ucsc.edu/index.php/GTEX2016