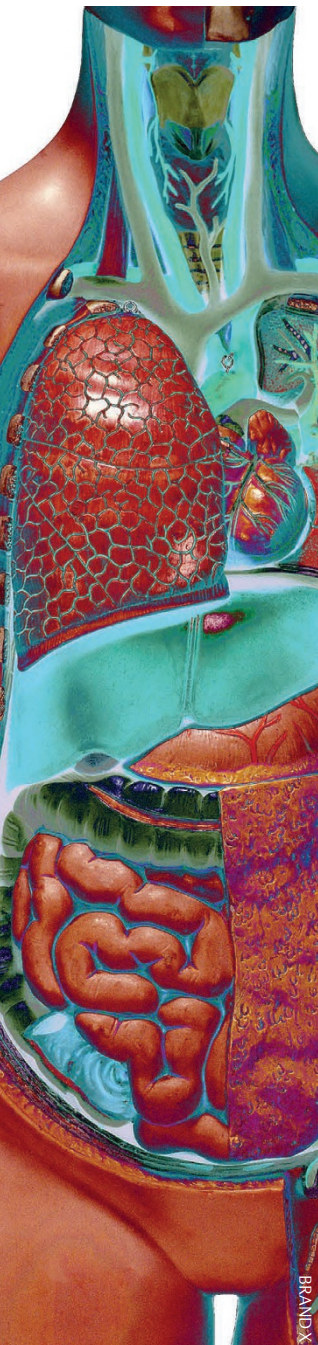### GENE EXPRESSION

Principles of gene regulation across tissues

The Genotype–Tissue Expression (GTEx) project is a large international consortium for investigating how the human genome encodes the principles of gene regulation across tissues. Following the results of the pilot phase of the project in 2015, several new papers describe the findings of the second, expanded phase of the project encompassing 7,051 samples from 449 donors. They report our most mature understanding to date of how genetic variants control gene regulation across human tissues.

The underlying strategy of the GTEx project is to profile a large collection of human donors post-mortem, so that all tissues are available for sampling. It involves genotyping each donor and profiling gene expression across 44 tissues. The genetic differences between donors can be cross-referenced with accompanying gene expression profiles to identify expression quantitative trait loci (eQTLs), that is, where DNA sequence variants are associated with altered gene expression.

The identification and characterization of two classes of eQTLs were the primary focuses of the main *Nature* paper authored by the GTEx Consortium. First were *cis*-eQTLs, in which a variant is associated with local expression changes (the transcription start site of the target gene is within 1 Mb of the variant). Although allele specificity was not part of the identification strategy, separate analyses at heterozygous sites showed that most *cis*-eQTLs exhibit allele-specific expression. Second were *trans*-eQTLs, in which variants are associated with long-range interchromosomal gene regulation.

There were notable differences between the two eQTL classes. Across the 44 tissues, *cis*-eQTLs could be identified for most genes (for 19,725 protein-coding genes), whereas *trans*-eQTLs were identified for only 93 genes, although the authors note that the study is relatively underpowered for detecting *trans*-eQTLs. Furthermore, *cis*-eQTLs were less tissue specific in their regulatory effects than *trans*-eQTLs. *Cis*-eQTLs were enriched for both promoter and enhancer elements, indicating direct regulatory effects on nearby genes. By contrast, *trans*-eQTLs often overlapped with *cis*-eQTLs and enhancers, consistent with the view that many *trans*-eQTLs act through local *cis* regulation of the expression of a protein (such as a transcription factor), which then mediates broader effects on gene expression interchromosomally.

As most genetic variants that have been associated with complex traits (including diseases) occur in non-coding regions, resources such as GTEx can help to link these variants to putative target genes and mechanisms in relevant tissues, especially as many of the GTEx tissues are inaccessible in living humans. Indeed, the authors note that approximately half of known complex-trait-associated loci colocalize with a GTEx eQTL.

Other co-published *Nature* papers mined this latest GTEx data set and focused on specific biological fields. Li *et al.* investigated the effects of rare genetic variants and found some that are associated with major gene expression alterations across tissues. They used this information to create their RNA-informed variant effect on regulation (RIVER) tool to facilitate the functional interpretation of rare

variants. Tukiainen *et al.* investigated X chromosome inactivation (XCI) and characterized heterogeneity in silencing of X chromosome genes across donors and tissues. Tan *et al.* studied RNA editing and deciphered RNA editing principles across genomic loci and tissues, including the identification of new regulators of RNA editing such as AIMP2.

In three accompanying GTEx papers in *Genome Research*, Mohammadi *et al.* quantified the effect size of *cis*-eQTL-associated gene expression changes, Saha *et al.* used gene co-expression networks to characterize tissue-specific regulation of transcription and splicing, and Yang *et al.* devised a bioinformatic tool for dissecting *trans*-eQTLs that are mediated through a local *cis*-eQTL effect. Finally, a commentary in *Nature Genetics* describes the Enhanced GTEx (eGTEx) project, which aims to provide a more complete mechanistic understanding of how genetics influences traits, by incorporating other omics data types such as epigenomic features, RNA modifications and proteomics.

Overall, these large, multilayered and continually maturing data sets are valuable resources for understanding how genetic variants influence physiological and pathological processes.

Darren J. Burgess

ORIGINAL ARTICLES The GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017) | Li, X. *et al.* The impact of rare variation on gene expression across tissues. *Nature* **550**, 239–243 (2017) | Tukiainen, T. *et al.* Landscape of X chromosome inactivation across human tissues. *Nature* **550**, 244–248 (2017) | Tan, M. H. *et al.* Dynamic landscape and regulation of RNA editing in mammals. *Nature* **550**, 249–254 (2017) | Mohammadi, P. *et al.* Quantifying the regulatory effect size of *cis*-acting genetic variation using allelic fold change. *Genome Res.* http://dx.doi.org/10.1101/gr.216747.116 (2017) | Saha, A. *et al.* Co-expression networks reveal the tissue-specific regulation of transcription and splicing. *Genome Res.* http://dx.doi.org/10.1101/gr.216721.116 (2017) | Yang, F. *et al.* Identifying *cis*-mediators for *trans*-eQTLs across many human tissues using genomic mediation analysis. *Genome Res.* http://dx.doi.org/10.1101/gr.216754.116 (2017) | eGTEx Project. Enhancing GTEx by bridging the gaps between genotype, gene expression, and disease. *Nat. Genet.* http://dx.doi.org/10.1038/ng.3969 (2017)