Nature Reviews Genetics | Published online 9 May 2016; doi:10.1038/nrg.2016.64

HUMAN GENETICS

Splicing: linking genetic variation and disease

PhotoDisc/Getty Images

It is well known that noncoding genetic variants contribute to complex traits and diseases, but the mechanisms through which they act are not fully understood. Now, a study published in *Science* reports that RNA splicing is a primary link between genetic variation and complex disease.

The authors set out to study the effects of genetic variants on different aspects of gene regulation, from chromatin to protein expression, using human lymphoblastoid cell lines (LCLs). To this end, they used a range of methodologies to quantify different regulatory traits including chromatin modifications, chromatin accessibility, DNA methylation, mRNA levels, transcription rate, RNA decay rate, translation levels and protein levels.

Then, using a variant-mapping pipeline and genomic data for LCLs, they mapped quantitative trait loci (QTLs) for the eight molecular phenotypes. To assess the regulatory levels at which QTLs act, they looked for overlap between expression QTLs (eQTLs) and QTLs associated with chromatin level traits, and determined that 65% of eQTLs exert their primary effect on expression at the chromatin level. The remaining 35% of eQTLs were enriched in transcribed regions, suggesting that many eQTLs act on transcriptional or post-transcriptional processes in a chromatin-independent manner.

To explore the effect of QTLs on splicing, the authors developed a novel tool called LeafCutter, which they used to identify 2,893 splicing QTLs (sQTLs). Whereas the majority of sQTLs had no effect on gene expression levels, 89% were found to directly affect predicted protein sequences. Notably, sQTLs were modestly enriched for association with chromatin-level phenotypes, providing direct evidence that splicing can be altered through changes at the chromatin level.

To assess the involvement of different regulatory mechanisms in complex traits and diseases, the authors searched for functional annotations among their data that were located in regions associated with rheumatoid arthritis, multiple sclerosis, Alzheimer disease, schizophrenia, height and body mass index. As expected, eQTLs were enriched in genome-wide association study (GWAS) data for diseases; however, sQTLs were enriched to a similar extent or in the case of multiple sclerosis to an even greater extent than eQTLs, indicating that variants that affect splicing make a major contribution to the genetics of complex diseases.

Together, these studies demonstrate that common genetic variation affects complex traits in three primary ways — through direct effects on expression, through effects on chromatin and through direct effects on splicing. Importantly, these studies reveal a major role for RNA splicing in linking genetic variation to complex traits and diseases, highlighting the importance of future investigations in this field.

Denise Waldron

ORIGINAL ARTICLE Li, Y. I. *et al.* RNA splicing is a primary link between genetic variation and disease. *Science* http://dx.doi.org/10.1126/science.aad9417 (2016)

WEB SITE

LeafCutter: https://github.com/davidaknowles/leafcutter