nature genetics

Pinpointing expression differences

Common gene variants influencing transcript levels can now be reproducibly identified by genome-wide screens. Some of the same variants contribute to clinical traits.

Genetic variants can affect the quantity and timing of transcription and transcript stability, as well as splice efficiency and the ratio of alternative splices, all of which can be detected as differential transcript levels by array methods. Three papers in this issue use linkage and association to identify variant loci that are correlated with RNA transcript abundance of human genes expressed in peripheral blood lymphocytes and immortalized lymphoblastoid cells.

Harald Göring and colleagues (p 1208) studied the loci influencing transcript levels in native peripheral blood lymphocytes by carrying out linkage analysis with 432 polymorphic microsatellites. One of these quantitative trait loci for expression (eQTLs) corresponds to a locus influencing HDL cholesterol levels. They found 42 eQTLs with 70% or greater heritability that they considered 'essentially monogenic' in their contribution to expression level of the closely linked transcription unit. However, the overall picture is polygenic. The mean effect size for eQTLs mapping to the same locus as the transcription unit at which RNA level was measured (*cis* eQTLs) was about 5% of the variance in transcript levels. There was a long tail, with over half of *cis* eQTLs accounting for more than 2% of the variance in transcript levels.

An earlier study by Morley *et al.* (*Nature* **430**, 743–747; 2004) identified *cis* regulators, which they defined as variants within 5 Mb of the measured transcription unit, and *trans* regulators that mapped outside this window. Göring *et al.*, using a stricter definition in which a *trans* eQTL had to map to a different chromosome from the target gene, were able to replicate 11 *cis*-acting loci of the Morley study but did not find 4 *trans*-acting loci reported by Morley *et al.* *Trans* effects are intrinsically harder to study: because there is no a priori hypothesis as to the location of a *trans*-regulatory element, detecting *trans* effects requires a very much larger number of tests of association or linkage. This means that small effects of *trans* variants are harder to detect than corresponding effects of *cis* variants.

Anna Dixon and colleagues (p 1202) report an eQTL associated with the same SNPs found in an association to childhood asthma recently reported by the group (Moffatt *et al.* *Nature* **448**, 470–474; 2007). This study also confirms that 'master regulators' of strong effect (at a distance >100 kb from the target transcript) are not widespread. Only 13 SNPs were associated with more than 10 heritable expression traits among some 6,660 genes with heritable expression differences (heritability, H^2, of >0.3). Interactions between SNPs regulating gene expression were demonstrated, providing an interesting area for future research.

Barbara Stranger and colleagues (p 1217) found 1,348 *cis* eQTLs and 180 *trans* eQTLs by association of 2.2 million SNPs in 270 individuals comprising the families used in the HapMap study. This study is notable for the number of associations that were replicated in multiple populations; indeed, 57 of the *cis* eQTLs and 5 of the *trans* eQTLs were found in all four populations examined. In 95–97% of the shared associations, the direction of the allelic effect was the same across populations. Interestingly, most of the *cis*-regulatory effects found map very close to the transcriptional start site of the target gene and were enriched in regions of high sequence conservation. In this study, *trans* regulation was defined as variants mapping >1 Mb from the center of the measured transcription unit. Because of the statistical implications of the large number of possible tests for *trans* eQTLs, Stranger and colleagues limited their examination to some 25,000 candidates—nonsynonymous SNPs and those predicted to affect splicing or miRNAs.

Despite the confidence engendered by the high degree of concordance among these studies, the differences reported are likely to be biologically interesting. Immortalized lymphoblasts that are clonal or of restricted lineage might be more suitable for the study of *trans* regulation and might be more readily studied without the environmental influences and transcriptome diversity found in a mixed lymphocyte population *in vivo*. However, as Stranger *et al.* discuss, *trans* regulation also involves intercellular communication in a multicellular organism and is consequently intrinsically harder to study. ■