Commentary

Incorporating Mouse Genome
Mammalian Genome
Genes and Phenotypes

The Genetics of Gene Expression

This special issue of Mammalian Genome highlights the merger of gene expression profiling with quantitative trait locus analysis. Inbred line crosses of model organisms provide a powerful tool for investigating complex trait genetics. Yet even the largest crosses provide only a coarse level of resolution, limiting our ability to identify genes. Transcriptional profiling provides a whole genome view of a primary step intermediate between the genetic variation encoded in DNA and the expression of this variation as a physiological trait. The realization that transcript levels can be analyzed as quantitative traits in a genetically segregating population has given rise to a synergistic experimental approach that promises to provide a mechanistic link between genetic and phenotypic variation.

A defining feature of transcript abundance as a quantitative trait is its correspondence with a precise genomic location – the structural gene and its surrounding regulatory elements. Expression QTL that map near the structural gene are abundant and are often observed to have large effects. These cis-QTL most likely represent polymorphisms in regulatory elements or in the gene itself. Equally interesting and abundant are the trans-QTL that map to locations distinct from the structural gene. The dramatic clustering of trans-QTL suggests that these genomic regions harbor polymorphisms that have widespread impact on the fundamental biology of the organism. They may represent a homeostatic response that provides robustness to changes in the internal environment. Whatever the explanation may be, the extensive networks of genes associated with trans-QTL attest to the dynamic and global nature of transcriptional regulation.

The genetic analysis of gene expression data poses some challenging computational problems. Of these, the scale up of genome scan analysis to include thousands of traits and the requisite multiple testing corrections are the most obvious. The real analytic challenges however must address the inverse problem of reconstructing the genetic architecture of transcriptional regulation based on the observed response of the transcriptome to perturbations. Knock-out models provide singular perturbations that are relatively easy to interpret but are limited in scope. Genetic crosses, on the other hand, generate a complex set of perturbations in which natural allelic variants at a multitude of loci are varied simultaneously. It remains an open problem to establish methods by which we can interpret these data as an integrated whole.

At this time, I do not believe that we fully understand either the scope or the limitations of this powerful new approach to genetic analysis. Gene expression in mammals is mediated by multiple transcription factors that interact through both cooperative and competitive binding to promoters. The potential for polymorphisms in both structural genes and their regulatory elements, suggests that pleiotropy and epistasis should be common features of the genetic architecture of gene expression. Moreover, many critical components of the biological system, such as metabolites, proteins and protein modification states, remain unobserved but the technologies required to measure these quantities are rapidly maturing. The articles in this issue can only hint at the potential for ''systems genetics'' to provide new and fundamenal insights into mammalian biology.

Gary A. Churchill

The Jackson Laboratory
Bar Harbor, Maine