Nature Reviews Genetics | AOP, published online 10 June 2008; doi:10.1038/nrg2403

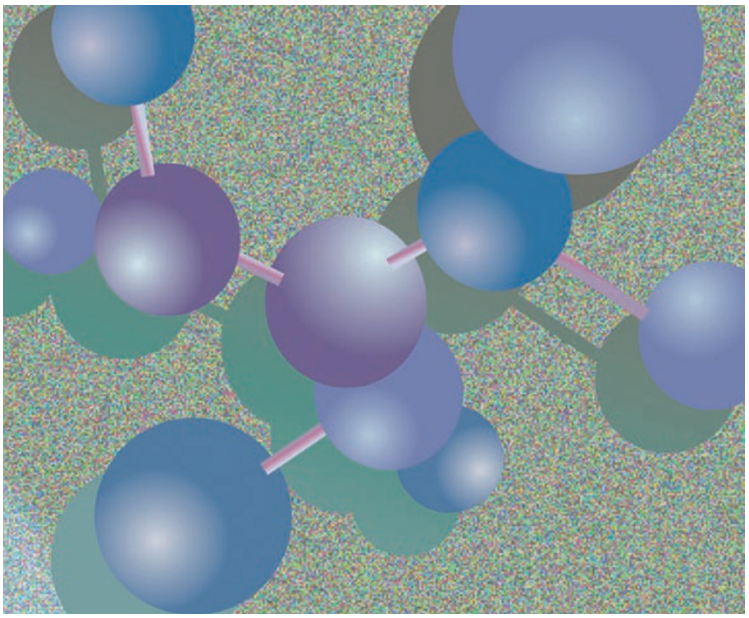
HUMAN DISEASE

Joining the dots from SNPs to proteins

A recent study provides the first genome-wide investigation of the relationship between common human genetic variation and levels of disease-related proteins. The findings have implications for understanding the aetiologies of complex diseases and how genetic variation influences the protein output of the genome.

Genome-wide association studies have made important inroads into understanding how common genetic variation influences complex disease. But the variants that have been identified generally have only small effects on the phenotypes of interest and often provide little insight into the functional basis of the association. A complementary strategy is to look for effects of genetic variation on gene expression. Although many variants that influence mRNA levels have been identified, the question remains of how such variation ultimately influences proteins levels, given the range of processes that affect protein abundance.

To begin answering this question, Melzer and colleagues looked for associations between almost half a million common SNPs and the blood levels of 42 proteins for which altered abundances in plasma or serum are correlated with disease status. The proteins chosen were relevant to a range of conditions, from metabolic and inflammatory disorders to HIV progression. Using samples from 1,200 individuals from the Italian InCHIANTI study, and applying stringent significance criteria, the authors identified 8 *cis* associations (that is, involving SNPs within 300 kb of the relevant gene). A single strong *trans* effect was also detected, but not replicated, possibly owing to differences in protein measurement between studies.



Some of the *cis* associations were already known: for example, between variation in the interleukin-6 receptor (IL6R) gene and levels of functional IL6 protein. Previously unreported associations were also uncovered, such as for the macrophage inflammatory protein beta (MIP-1-beta; also known as CCL4L), which is implicated in the development of AIDS. In each case, the effect of the variant on protein levels was large in comparison to the effect sizes known for variants that have been reproducibly associated with complex human traits. The authors also gained some insights into how the variants they identified might influence protein levels. These mechanisms include altered proteolytic processing, copy number variation and effects on transcript levels, although in most cases further studies are needed for confirmation.

These findings convey a clear message: common genetic variation has marked effects on protein levels, both through transcript abundance and other mechanisms. A fuller picture of this relationship will require further studies to detect a wider range of effect sizes for a larger number of proteins. The results also have direct implications for understanding disease mechanisms, for example, in determining whether altered levels of certain proteins are causes or consequences of disease.

Louisa Flintoft

ORIGINAL RESEARCH PAPER Melzer, D. *et al.* A genome-wide association study identifies protein quantitative trait loci (pQTLs). *PLoS Genet.* **4**, e1000072 (2008)
FURTHER READING McCarthy, M. I. *et al.* Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews Genet.* **9**, 356–369 (2008)