EPIGENETICS

Identical twins: epigenetics makes the difference

With their identical genetic make-up, monozygotic twins are firm favourites for sorting out genetic effects from those of the environment. But in a recent study it was the differences between identical twins that made them such fascinating research subjects: widespread epigenetic differences between twins that accumulate over the years indicate an important way that age and the environment might affect human health.

Manel Esteller and colleagues measured the total levels of two key epigenetic modifications — DNA methylation and histone acetylation — across the genomes of 40 pairs of monozygotic twins. For 65% of the pairs the twins had almost identical epigenetic profiles, but for the remaining 35% there were significant differences. Interestingly, there was a clear relationship between the age of the twins and the amount of difference between them. The amount of epigenetic difference was also correlated with spending large amounts of time apart and having different medical histories. So changes accumulated over time or influenced by environmental factors seem to have important effects on the epigenome.

To determine the biological significance of these epigenetic differences, the authors used a methylation fingerprinting technique in which distinct DNA bands correspond to individual methylated regions. Sequencing the bands that differed between identical twins revealed that although 52% of them corresponded to repetitive regions, the remainder reflected changes at predicted or known genes.

Importantly, the regions that were differentially methylated between twins also included CpG islands located in promoter regions, indicating potential effects on gene expression. This was confirmed by microarray analysis: whereas the expression profiles for pairs of 3-year-old twins were almost identical, there were large differences between the profiles of 50-year-old twins.

Alterations in gene expression that arise from global epigenetic changes over time are likely to have an important influence on susceptibility to many types of disease. The next challenge will be to work out how these changes arise: do they result from the cumulative effects of defects in epigenetic maintenance or transmission, or do environmental factors such as diet and exposure to pollutants have a role?

Louisa Flintoft

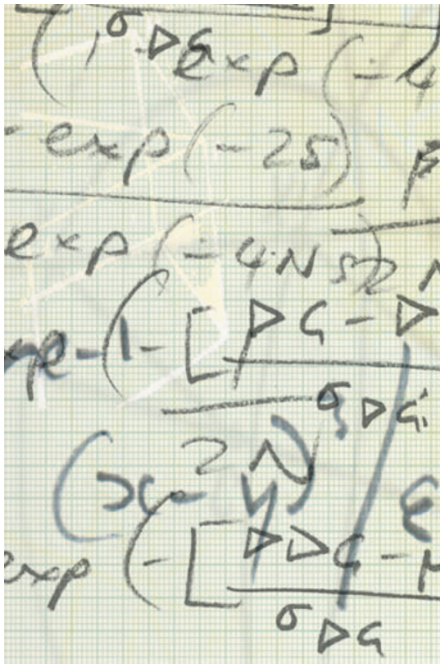
References and links

ORIGINAL RESEARCH PAPER Fraga, M. F. *et al.* Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl Acad. Sci.* **102**, 10604–10609 (2005)

FURTHER READING Fazzari, M. J. & Greally, J. M. Epigenomics: beyond CpG islands. *Nature Rev. Genet.* **5**, 446–455 (2004) | Robertson, K. D. DNA methylation and human disease. *Nature Rev. Genet.* **6**, 597–610 (2005)

GENE REGULATION

Copious calculations



Microarrays generate an abundance of gene-expression data, but finding the numerous loci that influence variation in expression poses huge computational and statistical challenges. Kruglyak and his colleagues have now published a method of linkage analysis that simultaneously maps the many QTLs that are responsible for variation in gene-expression 'traits'.

The authors applied their method to expression data from DNA microarrays in yeast. They crossed two haploid strains of *Saccharomyces cerevisiae*, and from the segregants obtained gene-expression measurements for more than 6,200 ORFs and found strain-specific genotypes at more than 3,300 loci. An established approach for identifying pairs of loci that affect the expression of a gene is 'two-dimensional' analysis, in which all pairs of markers are tested for linkage. But applying this approach to the yeast experiment would require more than 27 billion tests!

Instead, the authors developed a sequential approach, which begins by identifying the locus with the most significant effect on the gene-expression trait. Next, the locus that has the most significant effect on the trait over and above the effect of the first is identified; this sequential search can be continued for multiple loci. Locus-specific probabilities are combined to calculate the probability that all loci simultaneously affect the trait. These probabilities are used to estimate the false discovery rate among the set of significant traits (a significant trait is considered to be false positive if any of its selected loci are false).

By applying this method to the yeast expression data, the authors showed that their approach is more rigorous than the two-dimensional scan. They went on to use their identified loci to investigate the genetic architecture of gene-expression traits, showing that many traits are regulated by both *cis*-acting and *trans*-acting loci. They also found that several gene-expression traits share linkage to the same genomic position; so their method can also help to uncover regulatory networks.

Jenny Bangham

References and links

ORIGINAL RESEARCH PAPER Storey, J. D. *et al.* Multiple locus linkage analysis of genomewide expression in yeast. *PLoS Biol.* **3**, e367 (2005)