

## COMPARATIVE GENOMICS

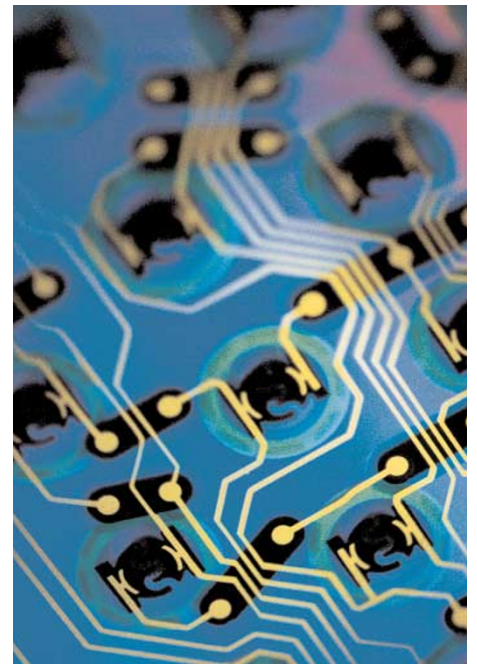
## Gene networks go global

Geneticists are used to thinking 'big' — whole-genome sequencing and genome-wide expression profiling are almost the order of the day. If the scale of these projects is no longer daunting, then translating the mountain of data they produce into something meaningful certainly is. In an intrepid and ambitious effort to get the most from genome-wide surveys, Sven Bergmann and colleagues have combined large-scale gene sequence and expression data for six evolutionarily distant organisms. The gene networks that emerged from this systematic, macro-evolutionary comparison of gene expression provide a model for understanding how the cell is built and designed and, on an immediately applicable level, provide a framework for gene annotation.

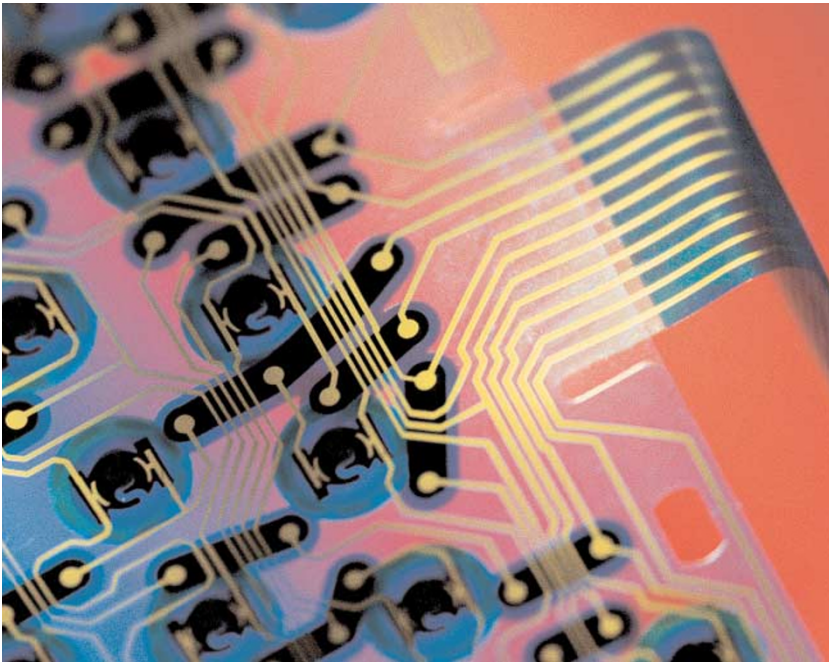
When it comes to gene sequence and expression, there is no shortage of available data, so the authors chose to take six 'post-genomic' organisms — bacteria, yeast, plant, nematode, fruitfly and human — and to exam-

ine the relationship between 40,000 genes using published sequence information and expression patterns. Step one was to determine whether co-expression of genes with similar function is conserved among species. Indeed, co-expression was conserved, particularly among genes that are involved in core (for example, metabolic) cellular functions. What's more, several clusters of co-expressed genes — 'transcriptional modules' — were conserved across the six species.

This allowed the authors to turn to the second issue: if transcriptional modules (and their components) are largely conserved among species, can the same be said of the relationship between them? Modules can be defined at various levels of stringency, but all gave the same result: although the expression of a few functionally-related modules was correlated across organisms (such as those for rRNA processing), the relationship between most of them was unique to a given species.



Probably the most innovative aspect of the work was to view gene expression networks from a global perspective: what do they look like and which properties can we infer from their topology? The expression data can be depicted as a tree of genes that are connected according to the degree of their co-expression. The picture of the tree that emerges from mathematical analysis is one that is dynamically evolving and that is rich in highly connected genes. Because these highly interconnected — so called 'hub' — genes are most likely to



be essential and evolutionarily conserved, they might correspond to those that were added at an early stage in the evolving network.

In summary, therefore, gene networks all look very similar despite differences in the behaviour of individual gene groups.

By combining extensive gene expression and sequence data, Bergmann *et al.* have taken comparative genome analysis to a higher, more powerful level. The use of expression data in comparative studies has its limitations, however, several

of which will fade if complete data sets are generated, as they almost certainly will.

*Tanita Casci*

#### **References and links**

**ORIGINAL RESEARCH PAPER** Bergmann, S. *et al.* Similarities and differences in genome-wide expression data of six organisms. *PLoS Biol.* 15 Dec 2003 (doi: 10.1371/journal.pbio.0020009)

**FURTHER READING** Stuart, J. M. *et al.* A gene-coexpression network for global discovery of conserved gene modules. *Science* **302**, 249–255 (2003)

#### **WEB SITES**

**Naama Barkai's laboratory:** <http://barkai-serv.weizmann.ac.il/GroupPage>

**Additional data and software:** <http://barkai-serv.weizmann.ac.il/ComparativeAnalysis>