COMPUTATIONAL BIOLOGY

A new algorithm weighs in

The computational approaches that are used to identify genetic regulatory networks have traditionally used information either from expression data sets or from genome-wide location analysis of DNA-binding regulators. Although the algorithms that analyse these different types of information have proven useful, their power is inherently limited by the fact that expression and binding data each provide only partial information — expression data provides only functional or indirect evidence and binding data only physical location information. Now, David Gifford and colleagues describe a new algorithm, GRAM (for genetic regulatory modules), which uncovers genetic regulatory networks by combining functional and physical information.

So, how does the GRAM algorithm work? First, it examines DNA-binding data and identifies sets of genes that are bound by common sets of transcriptional regulators. Next, it uses expression data to identify a subset of the genes, the expression of which is highly correlated. Finally, the algorithm searches the DNA-binding data again, using less stringent criteria, to find more genes of similar expression that are also bound by the same transcription factors.

The authors put the GRAM algorithm to the test in *Saccharomyces cerevisiae* by applying it to genome-wide location data for 106 transcription factors and >500 expression experiments — in this way, they identified 106 sets of co-expressed genes that are bound by the same transcription factors. These so-called gene modules contain 655 genes and are regulated by 68 transcription factors. Because gene modules link the common expression pattern of a set of genes to a set of regulators, the authors were also able to predict that 11 of the regulators were activators — a finding that was confirmed by previous experimental work.

Encouragingly, the GRAM algorithm generally placed genes that function in similar biological pathways in the same gene modules, and the regulators assigned to these modules were consistent with their known roles. In addition, for those gene modules controlled by more than one regulator, there was supporting evidence for a physical or functional interaction between the regulators.

That Gifford and colleagues were able to validate the results of the GRAM algorithm gave them confidence that it could be successfully used to analyse new data sources. Indeed, they generated a new genome-wide location analysis data set for the *S. cerevisiae* rapamycin response and were able to infer 39 new gene modules of 317 genes that were regulated by 13 transcription factors. The inferred network contained many established regulatory interactions, as well as some that were unexpected.

So, the GRAM algorithm successfully combines elements that take into account both mechanism (protein–DNA-binding events) and the consequences of mechanism (expression data) to provide a mechanistic explanation of genetic regulatory networks. This new algorithm is also "…particularly useful for uncovering how certain regulators may act in multiple biological pathways". And armed with the information generated by the GRAM algorithm, new interactions can be validated by directed experimental studies.

Natalie Wilson

References and links

ORIGINAL RESEARCH PAPER Bar-Joseph, Z. *et al.* Computational discovery of gene modules and regulatory networks. *Nature Biotechnol.* **21**, 1337–1342 (2003)

WEB SITES

David K. Gifford's laboratory: http://psrg.lcs.mit.edu/~gifford

Java implementation of the GRAM algorithm: http://www.psrg.lcs.mit.edu/GRAM/Index.html