Evolutionary change in the functional specificity of genes

Resource

INTERNET

Genome expression on the World Wide Web

The study of gene expression traditionally has been pursued through a combination of biochemical, genetic, and molecular biological studies. Genome sequences and new technologies have recently provided new approaches to study gene expression. By using high-density DNA microarrays or ‘DNA chips’, which consist of either oligonucleotides or cDNAs attached to a solid phase, researchers are now also able to measure the level of thousands of mRNAs simultaneously. This allows investigators to identify the set of genes influenced by a physiological event or a particular mutation and could ultimately allow biologists to understand the transcriptional programs of the cell.

Several groups have performed these kinds of chip experiments with the baker’s yeast, Saccharomyces cerevisiae, under a variety of conditions. Yeast currently has several advantages over meta-

omes for genome-wide expression studies. The entire yeast genome has been sequenced, so the expression level of every gene can be measured. The small size of the yeast genome, which consists of approximately 6200 genes, means that it takes fewer data points to provide complete information about the organism’s transcriptional state. Analysis of genome-wide expression data is more easily performed in the context of the substantial yeast literature. The genetic tractability of yeast permits efficient experimental examination of models that emerge from genome-wide expression data.

Web-based expression data

Papers featuring genome-wide expression experiments generally have accompanying web sites which are listed in Table 1. There are a few core features that are found in most of these sites. The ability to search a database by gene name allows users to track any gene of interest in a given experiment. Most experiments

References


5. Sulston, J.E. and Horvitz, H.R. (1983) Polyadenylated cell-


report expression values as a fold change from some standard experimental condition (e.g., wild type or time zero) relative to a condition of interest (e.g., mutant or time X). Users can search the database for those genes whose expression has changed a particular amount. When genes are listed, a brief annotation is supplied with the gene name which gives users some understanding of the function of genes with which they are unfamiliar.

Finally, most sites allow the user to download their data in tabular form. This is particularly useful for investigators who have devised their own methods to analyse and present data.

Brown and co-workers have used DNA chips to study various physiological processes in yeast by analysing the state of gene expression over a time course11. The user can query the data based on a minimum or maximum fold change for any number of time points and can therefore effectively retrieve the genes which exhibit a particular pattern of expression. The site accompanying Spellman et al.12 provides time course data for the cell cycle but displays it graphically for an easy snapshot of the transcriptional profile of a particular gene.

Other investigators have used genome-wide expression experiments to analyse various components of the transcription apparatus11,21. They use mutants in genes to determine what contribution various transcription factors make towards gene expression. The site supporting Holstege et al.12 allows the user to identify genes affected by the loss of a transcription factor and to list them according to functional categories. This provides insight into those transcription factors according to functional categories. This provides insight into those transcription factors and to list them towards gene expression. The site supporting Holstege et al.12 allows the user to identify genes affected by the loss of a transcription factor and to list them according to functional categories. This allows users to take a list of gene products and group them by function, metabolic pathway, or biochemical complex which will help uncover how a particular physiological process is regulated. It will be useful to determine and graphically display the intersection of the set of genes affected in one experiment with the set of genes affected in another; the experiment is to determine whether two cellular processes use similar regulatory mechanisms. Such analysis should help uncask the power of genome-wide expression experiments and reveal more of the transcriptional regulatory circuitry of the cell. Finally, it is particularly interesting to consider the development of a computational program capable of using large amounts of expression data to predict the transcriptional behavior of cells.

**Transcriptionome**

Two groups have described the yeast mRNA population in terms of the level of every detectable mRNA species; this population has been called the transcriptionome (Table 1). Velculescu et al.23 used serial analysis of gene expression (SAGE) to measure the number of copies of a given mRNA per cell. Holstege et al. used DNA chips. Both data sets are available on the web, searchable by gene name, and provide easy access to additional information about the gene of interest, either through the SGD24 for SAGE, or through YPD5 for the DNA chip data.

**Analysis**

As the study of gene expression through genome-wide analysis is new, the development of methods and tools for analysing expression data is in its infancy. Pfum et al.12 have created a program that allows investigators to perform cluster analysis on expression data. They use color bars to depict changes in gene expression and group genes which change in similar ways so that users can easily see genes whose expression is coordinately regulated. Several groups have written programs to search for over-represented sequence motifs in promoters10–12. These programs can be used in conjunction with genome-wide expression data to determine if a particular DNA sequence mediates the regulation of a group of genes. Future analytical programs will need to address several issues. The output from a given experiment is often a list of genes, many of which may be unfamiliar to the investigator. A program which allows users to take a list of gene products and group them by function, metabolic pathway, or biochemical complex will help uncover how a particular physiological process is regulated. It will be useful to determine and graphically display the intersection of the set of genes affected in one experiment with the set of genes affected in another; the experiment is to determine whether two cellular processes use similar regulatory mechanisms. Such analysis should help uncask the power of genome-wide expression experiments and reveal more of the transcriptional regulatory circuitry of the cell. Finally, it is particularly interesting to consider the development of a computational program capable of using large amounts of expression data to predict the transcriptional behavior of cells.

**Acknowledgements**

We thank Fran Lewitter for helpful comments. E.G.J. is a predoctoral fellow of the Howard Hughes Medical Institute.

---

**TABLE 1. Genome expression web resources**

<table>
<thead>
<tr>
<th>URL</th>
<th>Description</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.scripps.edu/GENEX/">http://www.scripps.edu/GENEX/</a></td>
<td>Serial analysis of gene expression (SAGE) to measure the number of copies of a given mRNA per cell.</td>
<td>6</td>
</tr>
<tr>
<td><a href="http://www.sanger.ac.uk/GENEX/">http://www.sanger.ac.uk/GENEX/</a></td>
<td>Yeast Proteome Database (YPD)</td>
<td>8</td>
</tr>
<tr>
<td><a href="http://www.mips.biochem.mpg.de/proj/yeast/transcription/mig1_contr.html">http://www.mips.biochem.mpg.de/proj/yeast/transcription/mig1_contr.html</a></td>
<td>MIG1</td>
<td>16</td>
</tr>
<tr>
<td><a href="http://www.mips.biochem.mpg.de/proj/yeast/browsing.html">http://www.mips.biochem.mpg.de/proj/yeast/browsing.html</a></td>
<td>Yeast Proteome Database (YPD)</td>
<td>8</td>
</tr>
<tr>
<td><a href="http://www.sanger.ac.uk/GENEX/">http://www.sanger.ac.uk/GENEX/</a></td>
<td>Yeast Proteome Database (YPD)</td>
<td>8</td>
</tr>
<tr>
<td><a href="http://www.mips.biochem.mpg.de/proj/yeast/explore.html">http://www.mips.biochem.mpg.de/proj/yeast/explore.html</a></td>
<td>MIG1</td>
<td>16</td>
</tr>
<tr>
<td><a href="http://www.mips.biochem.mpg.de/proj/yeast/browsing.html">http://www.mips.biochem.mpg.de/proj/yeast/browsing.html</a></td>
<td>Yeast Proteome Database (YPD)</td>
<td>8</td>
</tr>
<tr>
<td><a href="http://www.mips.biochem.mpg.de/proj/yeast/explore.html">http://www.mips.biochem.mpg.de/proj/yeast/explore.html</a></td>
<td>MIG1</td>
<td>16</td>
</tr>
</tbody>
</table>

---

**References**

Deification of the genes

The Genetic Gods. Evolution and Belief in Human Affairs
by J.C. Baley
Harvard University Press, 1998. $29.95 hbk (viii + 279 pages) ISBN 0 674 34625 4

Who or what are the ‘genetic gods’? According to the author, ‘they mastermind our lives, influencing our physical appearance, health, behavior, even our fears and aspirations. They constitute our maternal reason for being – for eating and sleeping, working and loving, baring and caring, forging relationships for procreation...we are their tickets to immortality...They are not gods, but our genes...Genes have special powers over human lives and affairs...genes exert influence over the course of nature...gene lineages are potentially immortal.’

Why does the author wish to equate genes with gods? One reason, or at least outcome, might be that it permits him to use terminology that is descriptive of the behavior of a god or gods (‘malevolent’, ‘magnanimous’, ‘self-serving’) when talking about genes, thereby engaging the attention of readers susceptible to the use of anthropomorphic (or perhaps deitomorphic) terms when scientific concepts are being described. However, the real reason, I believe, is far more serious and is revealed in the following quotations:

‘The emergence of Homo sapiens under natural evolutionary processes can be interpreted as an even more miraculous and awe-inspiring than human creation by a god.’

‘Human beings, like all other species on earth, are biological products of evolutionary processes, and as such are physical expressions of the genes, the “genetic gods”. Genes and the mechanistic evolutionary forces that have sculpted them thus assume many of the roles in human affairs traditionally reserved for supernatural deities.’

Regardless of what they are called, genes are tangible entities, with profound influences on humanity. Indeed, over the last century, the genetic gods would seem to have wrested from the supernatural gods considerable authority over human affairs. Does any room remain for a metaphysical god?

‘What is to be gained by an awareness of genetic operations and evolutionary processes when such knowledge challenges our faith in a loving and interventionist god?’

The thrust of the argument, therefore, is that it is not necessary to invoke religious beliefs to understand how mankind has reached its present state, not only in physical and functional terms, but also in terms of its social organization and, indeed, its moral and ethical precepts. Thus, ‘the richly