GENOMICS

Six degrees of separation

A new resource allows researchers to match the gene-expression signature of their system of interest to that of well-characterized chemical compounds—a hypothesis-generation tool with a bright future.

The Connectivity Map, put together by scientists at the Broad Institute of MIT and Harvard and recently described in *Science* (Lamb *et al.*, 2006), is a wealth of genomics information that bench researchers will find immediately useful.

This resource contains a collection of gene-expression profiles captured after treatment of cell lines with well-characterized small molecules—164 in the first release—with a broad range of activities. A suite of informatics tools allows users to compare a gene-expression profile of interest to those present in the database. The power of this type of analysis is illustrated in companion papers that appeared in *Cancer Cell* (Hieronymus *et al.*, 2006; Wei *et al.*, 2006).

The Connectivity Map was conceived with general users in mind. "It was very much a guiding principle," says Justin Lamb who led the Connectivity Map effort to "make it easy and accessible for people who might not necessarily be experts in genome analysis."

To appreciate how this new tool can help bench researchers, take, for example, someone who has screened a small-molecule library and found a compound causing a phenotype of interest. "Under the best circumstances," says Lamb, "you will be presented with a compound, a structure and a name; you've got some stuff in a tube that does something but it doesn't enlighten your research in any way." Enter the Connectivity Map, which might move things along considerably by providing functional annotation.

The investigator would upload a gene-expression signature for the compound to the Connectivity Map, which—using the ranking algorithms—would generate a list of compounds with the most similar gene-expression signatures, and thus likely to have the same activity. "The advantage for the bench researcher is that an experiment is immediately suggested," explains Lamb. It would then be fairly straightforward to figure out whether the compound of interest has the activity suggested by the Connectivity Map. An example of this ability to infer functionality by connecting the effect of an unknown

CELL BIOLOGY

SYSTEMS BIOLOGY FOR BEGINNERS

A web and print Focus on systems biology from Nature Publishing Group provides a practical introduction to a field that for all its promise still has many skeptics.

Pick an experimental biologist at random and ask their opinion on systems biology, and it is unlikely that you will receive an indifferent response. This burgeoning field with its massive data sets has boosters and detractors aplenty. A probable reason for this is the infancy of a field that as yet only rarely lives up to its promise of providing understanding at a more holistic level, inaccessible to classical research. Another reason may be its reliance on mathematical and computational approaches, which often remain intractable to bench-trained molecular biologists. Researchers new to systems biology can be put off by the combination of inaccessible tools and the present lack of contributions to biological understanding. The complexity of biology has made it clear however, that a new approach is needed.

In *Systems biology: a user's guide*, a joint web and print Focus from *Nature Reviews Molecular Cell Biology* and *Nature Cell Biology*, systems biology experts focus on explaining the methodologies behind systems-biology approaches for molecular cell biology. This information is aimed at investigators who want to know what systems biology is really about and whether it can enhance their own research. The collection of nine articles covers biophysical, statistical and physicochemical modeling approaches, and provides an overview of requirements for modeling-friendly data.

The most basic requirement for entering the field is an understanding of large databases and their use for data mining. The user's guide provides a list of important bioinformatics

small molecule to that of well-characterized compounds is presented in the study by Hieronymus *et al.*

Another application consists of making a connection between a disease state and the activity of small molecules present in the database as demonstrated by Wei *et al.* Using the Connectivity Map, this group was able to match the gene-expression signature of a drug-sensitive leukemia to that of rapamycin. When such an approach is successful, it immediately generates a testable hypothesis about pathways involved in the disease state. Again the benefit is in the shortcut to experimentation. Instead of staring at lists of up- and downregulated genes directly or indirectly involved, the Connectivity Map provides you with a very tangible chemical activity. As Lamb puts it, "It circumvents the lists of genes to go from disease to small molecule." In addition to a new hypothesis to test, defining a chemical activity involved also brings you much closer to therapeutic intervention.

These examples build a very good case for the usefulness of the Connectivity Map, but not all applications will be as successful, and using this resource will require some judgment calls. The results of a search come with a statistical qualifier, but there is no clear threshold yet for a real match, and this threshold will likely depend on the question that is asked. Finding connections between the activities of two chemical compounds may be relatively straightforward, whereas connecting the activity of a small molecule to a complex disease will not be as clearcut.

Although the Connectivity Map has already proven useful, Lamb and his team have plans for improvements. Of high priority is expansion of the database to include the 1,400 US Food and Drug Administration–approved drugs. This well-characterized set of compounds recapitulates a lot of different chemical activities. As Lamb points out, "It is the distillation of hundreds of years of medicinal chemistry." The team also appears eager to receive feedback from users to plan further developments.

Veronique Kiermer

RESEARCH PAPERS

Hieronymus, H., *et al.* Gene expression signature–based chemical genomic prediction identifies a novel class of HSP90 pathway modulators. *Cancer Cell* **10**, 321–330 (2006).

Lamb, J., *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).

Wei, G., *et al.* Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. *Cancer Cell* **10**, 331–342 (2006).

WEB SITES

The Connectivity Map homepage: http://www.broad.mit.edu/cmap/

resources, and reviews their design and use to help prospective users navigate this information trove. Data are of little use for modeling purposes if the information is not accurately defined, quantitative and reproducible; and it is important for prospective users to understand and implement data gathering and storage.

Regardless of the quality and annotation of the data, proper analysis will still be hindered by the incomplete nature of the data and experimental noise. Mathematical models must be derived from appropriate data and properly validated. The many options available for creating models can be daunting and the overviews of some of the key approaches provide guidance on what suits a given field best.

The goal of this user's guide is to help non–systems biologists better understand this important new dimension to biological investigation and prompt them to consider using these methods for their own data. For those who wish to explore the field in more depth the online guide provides a list of textbooks, and for those who want to explore collaborations or job opportunities there is a list of institutes specializing in systems biology. Real understanding of complex biological processes will continue to elude us if we rely solely on reductionist and largely intuitive approaches. Biologists have the tools to go farther, they only lack the data.

Daniel Evanko

WEB SITES

The *Systems biology: a user's guide* Focus homepage: http://www.nature.com/Focus/systemsbiologyuserguide/