SCIENCE@BERKELEY LAB

February 15, 2008

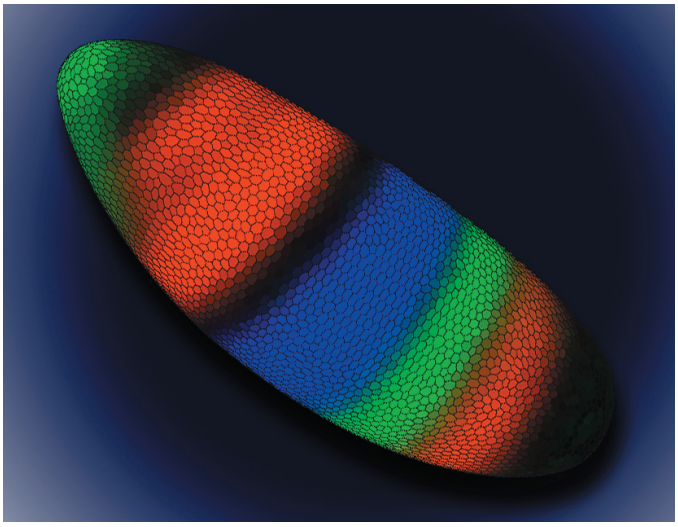
Mysteries of Genome Regulation

Gene regulators may bind promiscuously, but they often do nothing

Contact: Lynn Yarris, lcyarris@lbl.gov

Biologists are developing ever more sophisticated means to characterize molecular interactions in living systems, but a study by Berkeley Lab researchers suggests that many of the interactions being detected are functionally irrelevant. Their findings show that the transcription factors that choreograph early development in the fruit fly *Drosophila melanogaster* bind to a surprisingly wide array of genes—but that much of this binding has no effect on gene expression.

"As we develop new techniques that allow us to observe the molecular events in living systems at ever finer scale and increasing resolution, there is a natural tendency to assume that if something happens in a cell it must be important," says Michael Eisen, a geneticist with Berkeley Lab and the University of California at Berkeley. "But our results show that this assumption is not always correct, and this should be instructive for anyone working in any area of functional genomics."



This image shows a Drosophila embryo colored to show the expression patterns of early gene regulators. Each color represents the level of expression of one of three gene regulators, Knirps (green), Kruppel (blue), and Giant (red). Color intensity reflects a higher level of expression. The darker areas of the embryo are cells where none of these gene regulators are expressed, and the yellowish areas indicate that both Knirps and Giant are being expressed.

Eisen directed the study along with Mark Biggin, a biochemist with Berkeley Lab's Genomics Division. Their results are described in a paper published today in the open-access journal *PLoS Biology*. Eisen and Biggin were joined by Xiao-Yong Li and Stewart MacArthur as lead authors of the paper, which is freely available online.

The fruit fly is the preeminent model for the study of morphogenesis, the process by which embryonic cells are able to multiply and form three-dimensional arrays that eventually become tissues, organs, and finally entire organisms. The blueprints for this transformation are encoded in the DNA of every cell, and are "read out" by vast networks of transcription factors that regulate where and when genes are expressed.

To understand how these gene regulators coordinate morphogenesis, scientists must first learn the full range of genes under their direct control. To do this, the dominant scientific view holds that it is only necessary to identify the number and types of genes to which each regulator is bound.

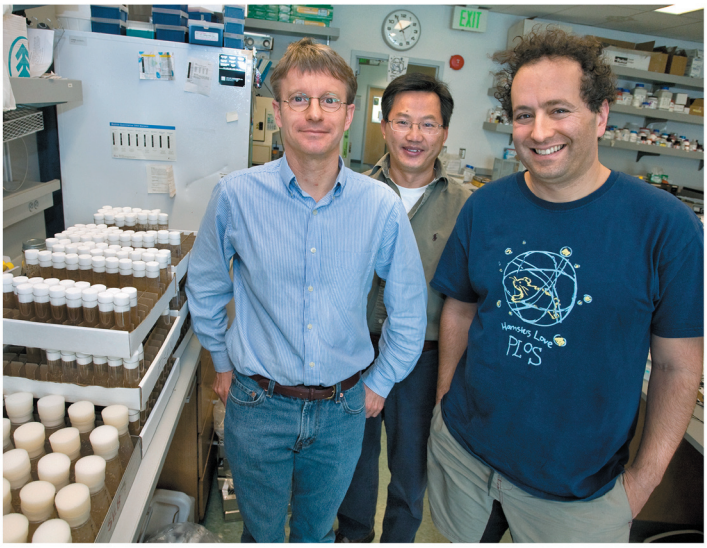
The Berkeley team set out to do just that using a popular technique known as ChIP/chip (for "chromatin immunoprecipitation analyzed on DNA microchips") to catalog the genomic locations bound by six transcription factors that shape patterning along the *Drosophila* anterior-posterior (head to tail) axis.

continued

While the set of known targets for each of these extensively studied transcription factors numbered in the dozens, the Berkeley researchers eventually found thousands of genomic locations reproducibly bound by each factor.

"This is several orders of magnitude more genes than these factors are thought to regulate," said Eisen. "This raises the question of what the function of the binding is."

Rather than simply classify regions as bound or unbound, as other researchers have done, the Berkeley team examined the full scope of binding observed in their ChIP/chip data and focused on quantitative differences in the amount of transcription factor bound to each gene. This perspective was motivated by pre-genome technology experiments led by Biggin in the 1990s, which showed that transcription factors bind to many more genes than expected, in what Biggin characterized as a "broad quantitative continuum".



(From left) Mark Biggin, Xiao-Yong Li and Michael Eisen led the Berkeley study that found that many of the interactions between gene-regulating transcription factors and DNA have no function.

(Photo Roy Kaltschmidt, Creative Services Office)

In the current study the Berkeley researchers found a clear relationship between the number of factor molecules bound at a given site and the site's role in gene regulation. Sites bound by the highest densities of molecules were generally associated with the patterning of a limited set of approximately 100 genes crucial to the developmental process. Many of the genes bound at intermediate levels, while not typically thought to have a role in development, were also regulated by these factors.

But much of the low-level binding detected at thousands of genes—while clearly representing real molecular interactions—appears to play no role whatsoever in regulating gene expression.

In the recent rush to survey the landscape of genome-wide, transcription-factor interactions with DNA (several dozen papers on the topic have been published in the last few years), there has been scant acknowledgement of the possibility that much of the binding observed in these studies may be nonfunctional. Based on their observations, the Berkeley Lab researchers argue that many of the regulatory connections proposed in earlier studies are likely to be incorrect.

"Realizing that much of the binding detected in recent genome-wide assays may be nonfunctional significantly impacts how the results of these experiments should be interpreted," says Biggin. "The analysis and conclusions of published ChIP/chip studies should be reexamined with this possibility in mind."

This work was carried out as part of a broader collaboration by the Berkeley Drosophila Transcription Network Project (BDTNP). The in vivo binding data and computational analyses were funded by grants from the U.S. National Institutes of Health.

continued

Additional information

"Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm," by Xiao-yong Li, Stewart MacArthur, Richard Bourgon, David Nix, Daniel A. Pollard, Venky N. Iyer, Aaron Hechmer, Lisa Simirenko, Mark Stapleton, Cris L. Luengo Hendriks, Hou Cheng Chu, Nobuo Ogawa, William Inwood, Victor Sementchenko, Amy Beaton, Richard Weiszmann, Susan E. Celniker, David W. Knowles, Tom Gingeras, Terence P. Speed, Michael B. Eisen, Mark D. Biggin, can be freely viewed and downloaded at http://biology.plosjournals.org/perlserv/?request=get-document&doi=10.1371%2Fjournal.pbio.0060027

The *PLoS Biology* website is http://www.plosbiology.org

More about Michael Eisen's research is at http://rana.lbl.gov/

More about Mark Biggin's research is at http://gsd.jgi-psf.org/biggin.shtml

More about the Berkeley Drosophila Transcription Network Project is at http://bdtnp.lbl.gov/Fly-Net/