

GENOMICS

Good enhancer hunting

A new strategy allows identification of enhancers in the human genome on a large scale.

In many respects, the sequencing of the human genome represents just the beginning of what will be a long and difficult, but exciting, journey for researchers. With the sequence of the human genome in hand, biologists are now faced with the rather daunting challenge of identifying and understanding the myriad functional elements it contains. Much progress has already been made in pinpointing the locations of protein-coding genes within the genome, but other functional elements—such as the DNA sequences that control when and where those genes are actively transcribed—remain elusive. Distal enhancers are one category of functional genomic elements that regulate gene expression. Locating these enhancers in the vast expanses of the genome can be difficult, however, especially because enhancers are frequently located quite far away from the genes that they regulate.

Recently, a team of scientists led by Len Pennacchio at Lawrence Berkeley National

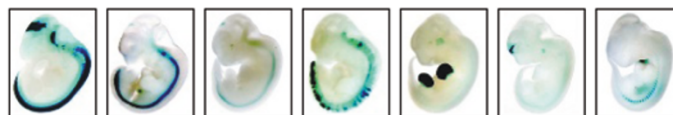


Figure 1 | Enhancers were identified based on their ability to turn on expression of a reporter gene (blue) in specific regions of a developing mouse embryo. Reprinted with permission from *Nature*.

Laboratory used an ambitious strategy to identify and characterize previously unknown enhancers in the human genome on a large scale. First, they performed DNA sequence comparisons to identify DNA sequences that were highly conserved between humans and other animals; they then experimentally tested these sequences for enhancer activity using an *in vivo* assay. It was not clear at the outset how effective this strategy would be; as Pennacchio explains, “Despite anecdotal evidence that previously identified enhancers appear highly conserved across species, we were unclear how well conservation as a starting point would yield active enhancers. This is due to the fact that there are so many different types of noncoding functions beyond transcriptional enhancement.”

Pennacchio and colleagues began their study by performing extensive sequence alignments between human genomic DNA sequence and genomic DNA sequence from mouse, rat and pufferfish. From these sequence alignments they were able to identify noncoding human genomic sequences that have been highly conserved over the course of evolution. High conservation suggests that a DNA sequence may have an important function, thereby constraining the sequence from changing very much. The researchers chose 167 of these highly conserved sequences to test whether they actually function as enhancers. For their *in vivo* assay, they generated transgenic mice carrying the human version of the highly conserved sequences. Each transgenic mouse

SYSTEMS BIOLOGY

ANNOTATING THE UNANNOTATED

By combining activity-based proteomics and metabolomics, researchers have developed a new systems biology strategy for characterizing enzymes in the context of metabolic networks.

By harnessing the power of mass spectrometry, researchers have been generating massive amounts of peptide and protein sequence data from just about every kind of biological sample. Although the databases containing these lists are certainly valuable resources, the effort to follow up with detailed studies to assign all of these proteins in biological contexts has lagged. Even in the very best-annotated protein classes, perhaps only 50% or so of the known members have been characterized.

But some researchers, such as Benjamin Cravatt at the Scripps Research Institute, find this part of the process the most exciting. Cravatt’s group has been at the forefront of efforts to develop activity-based protein profiling (ABPP), or functional proteomics, which they use to identify biologically interesting enzymes, especially those relevant to human diseases such as cancer.

Though the ABPP method allows them to discover enzymes in various functional classes, characterization requires steps for further analysis. Cravatt explains, “Our next goal was to

come up with an unbiased small-molecule profiling platform that would, in response to inactivating an enzyme, allow us to see what the metabolic consequences were in the broadest possible sense and hopefully link [the enzyme] into the metabolic network that it regulates.” For their first demonstration of this strategy, they chose the target KIAA1363, an enzyme that is found in particularly high levels in aggressive cancer cells, but whose biological context was uncharacterized.

Using competitive ABPP, Cravatt and his colleagues identified a small-molecule inhibitor of KIAA1363, named AS115, that strongly and selectively inhibited the enzyme’s activity. They treated several cancer cell lines with AS115 and compared the liquid chromatography–mass spectrometry metabolic profiles of the inhibitor-treated cells to those of untreated cells. The inhibitor treatment resulted in a reduction in the level of an unusual class of lipids, the monoalkylglycerol ethers, which were shown to be direct products of a KIAA1363-catalyzed reaction. Amounts of a secondary metabolite class, the lysophospholipids, also were reduced. In parallel, by performing short hairpin RNA-mediated knockdown of KIAA1363 (shKIAA1363), the researchers observed similar reductions in the same metabolite classes, but the downstream effects were even more dramatic

NEWS IN BRIEF

line carried a different DNA sequence, and the researchers engineered the mice such that DNA sequences that were *bona fide* enhancers could be identified on the basis of their ability to turn on expression of a reporter gene (Fig. 1).

Using this strategy, the researchers identified many new enhancers, with approximately half of the sequences they tested in mice functioning as actual enhancers. Pennacchio cautions, though, that this strategy does have certain limitations, and may be biased toward identifying particular classes of enhancers: “We have currently focused on an extreme version of human genome noncoding conservation and these elements are not randomly distributed. Rather they are highly clustered and biased toward transcription factors and other key developmental genes (likely due to the extreme constraint on the regulation of genes important in vertebrate body plan development)”. Additionally, enhancers that are active in an adult, but not during embryonic development, would be missed because the enhancer activity screen is only performed at an embryonic time point.

Nonetheless, this approach has proven to be very effective at enhancer identification. In fact, since their initial study, the researchers have continued their efforts to identify new enhancers and are doing so at an impressive rate. Pennacchio explains, “In this study, in our first pass, we tested ~170 elements in about a year and are now testing closer to 500 per year. With an enhancer identification success rate of ~50%, we are quickly surpassing the cumulative number of enhancers identified by all investigators to date.”

Jesse Potash

RESEARCH PAPERS

Pennacchio, L.A. *et al.* *In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* **444**, 499–502 (2006).

than with small-molecule inhibition. “The advantage of the small-molecule inhibitor is that you can get a quick picture of what the acute effects are of inactivating an enzyme,” explains Cravatt, whereas “the shRNA approach gives you a steady-state picture of what an inactivated enzyme looks like over many days,” which will certainly be important for evaluating potential drug targets. Additionally, mice injected with shKIAA1363 cancer cells showed reduced tumor growth compared to those injected with control cancer cells.

This powerful approach allowed the researchers to place KIAA1363 at a central node in an ether lipid signaling network. “This paper really shows, I think for the first time, that one can go into a complex biological model like a human cancer cell and annotate a disease-relevant protein using purely systems biology methods. We didn’t have to recombinantly express or purify the enzyme at all to figure any of this stuff out,” says Cravatt. As opposed to the more classical ‘test-tube’ enzyme characterization, Cravatt explains, “being able to do these experiments in living systems is really valuable because it allows you to circumvent potential *in vitro* artifacts related to what an enzyme can do versus what it really does do.”

Allison Doerr

RESEARCH PAPERS

Chiang, K.P. *et al.* An enzyme that regulates ether lipid signaling pathways in cancer annotated by multidimensional profiling. *Chem. Biol.* **13**, 1041–1050 (2006).

IMAGING AND VISUALIZATION**Welcoming an improved blue fluorescent protein**

Though a rainbow of genetically encodable fluorescent proteins are available for biological imaging applications, existing blue fluorescent protein (BFP) variants were rather dim, had low quantum yields and underwent rapid photobleaching. By repacking the BFP core to make the chromophore more rigid, Mena *et al.* have developed a new BFP named Azurite with greatly improved properties for biological imaging in both bacterial and mammalian systems.

Mena, M.A. *et al.* *Nat. Biotechnol.*; published online 19 November 2006.

PROTEOMICS**A quantitative picture of the synaptic vesicle**

Using a battery of proteomics, biophysical, imaging and modeling technologies, Takamori *et al.* have constructed the first detailed, quantitative model of a trafficking organelle—the synaptic vesicle. They determined its size, density and mass, its full protein and lipid composition, and the copy numbers of major constituents. This complete model sets the stage for future quantitative trafficking studies.

Takamori, S. *et al.* *Cell* **127**, 831–846 (2006).

STEM CELLS**Small-molecule stem-cell maintenance**

There is currently much to learn about the molecular cues that cause embryonic stem cells (ESCs) to remain in an undifferentiated state. Normally, when ESCs are grown *in vitro*, feeder cells are used to help them remain undifferentiated. Using high-throughput screening of chemical libraries, Chen *et al.* identified a new small molecule that can be used to maintain mouse ESCs in an undifferentiated state, even in the absence of feeder cells.

Chen, S. *et al.* *Proc. Natl. Acad. Sci. USA* **103**, 17266–17271 (2006).

SPECTROSCOPY**Elementary insights into an enzyme mechanism**

Using a combination of colloidal probe atomic force microscopy and a quartz crystal microbalance, Suzuki *et al.* obtained an extremely detailed picture of the specific elementary interactions taking place in an enzymatic reaction. For the first time, they were able to investigate the complex interactions between the heptaprenyl disphosphate synthase enzyme subunits, a cofactor and a substrate.

Suzuki, T. *et al.* *J. Am. Chem. Soc.* **128**, 15209–15214 (2006).

PROTEIN BIOCHEMISTRY**A method for selective sulfation**

Tyrosine sulfation is a common post-translational modification found in eukaryotic proteins, yet tools to probe the biological function of this modification have been lacking. Liu and Schultz now describe a method for the selective, recombinant incorporation of sulfotyrosines into proteins by an orthogonal tRNA–aminoacyl-tRNA synthetase pair assigned to the amber nonsense codon. They expressed sulfo-hirudin in *Escherichia coli*, which previously could not be made using recombinant methods. Liu, C.C. & Schultz, P.G. *Nat. Biotechnol.* **24**, 1436–1440 (2006).