Deciphering Development: Quantifying Gene Expression through Imaging

Author: PHILLIPS, MELISSA LEE

Source: BioScience, 57(8) : 648-652

Published By: American Institute of Biological Sciences

URL: https://doi.org/10.1641/B570803

Deciphering Development: Quantifying Gene Expression through Imaging

MELISSA LEE PHILLIPS

Scientists can now visualize developmental gene expression quantitatively in three dimensions and at single-cell resolution. Recent advances in optical microscopy and fluorescent genetic tags allow imaging of gene expression in live animals, as well. Eventually, researchers hope to construct virtual atlases of animal development.

For more than a hundred years, developmental biologists have used microscopic imaging to investigate embryonic development. By staining animal tissue at different developmental stages, they have seen how cells, tissues, and organs develop. More recently, scientists have created molecular tags that permit imaging of gene or protein expression. But these methods usually reveal little, if any, quantitative information. "The great majority of descriptions of gene expression patterns and morphology are qualitative," says Mark Biggin of Lawrence Berkeley National Laboratory in Berkeley, California.

A major problem with a qualitative approach to studying gene expression is that the human eye fails to see small, but potentially important, changes, Biggin says. "If you look with your eye, you can't see a small, twofold change in expression that occurs slowly and steadily over a whole series of cells." Also, precise quantitation is necessary for scientists to model computationally the intricate patterns of animals' developmental gene expression and morphology.

This zebrafish embryonic eye, magnified here 40 times, is 26 hours postfertilization. Green fluorescence marks cell nuclei, and red fluorescence labels cell membranes. The lens is the circular structure in the middle, and the retina is the ring around it. Image: Sean Megason, California Institute of Technology.

Recent advances in optical microscopy and fluorescent genetic tags are now allowing scientists to measure developmental gene expression not only quantitatively but also in three dimensions and at single-cell resolution. Some of these techniques also allow imaging in live animals. "We want to start with the egg, start with one cell, and then see how that cell divides, how the cells move around, and how an organism is basically constructed, just by watching it," says Sean Megason of the California Institute of Technology (Caltech) in Pasadena, California.

A key advance in this type of imaging came in 1992, when researchers first cloned and sequenced green fluorescent protein (GFP) from the bioluminescent jellyfish *Aequorea victoria.* In 1994, researchers reported that they had introduced GFP into the *Caenorhabditis elegans* genome and that they could use

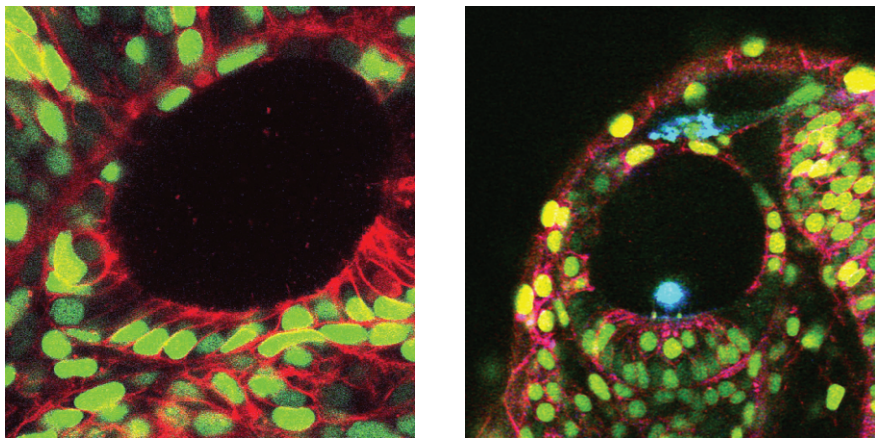
its fluorescence as a genetic marker in living tissues. Unlike some other labeling methods, GFP provides single-cell resolution of expression patterns and can be used in live, intact animals. Advances in microscopy came next, with the invention of laser-scanning microscopy (LSM). Unlike traditional light microscopy, LSM can generate three-dimensional (3-D) images by capturing a stack of two-dimensional (2-D) images and reconstructing them.

The combination of LSM and fluorescent tags now allows researchers to image cell movements and divisions, as well as the genes and proteins that each of these cells expresses during embryogenesis. Image-analysis algorithms then extract information about cell behavior and gene expression levels. "We want to be able to actually quantitatively measure that process," says Megason. "The imaging just serves as a way to digitize the information."

Eventually, researchers hope to construct virtual atlases of animal development. These atlases would map gene expression in each cell of an animal as its embryo develops. "Ultimately, you would want to have all genes throughout embryogenesis in all cells," Biggin says. Early studies have already revealed aspects of embryonic development that couldn't be seen with 2-D, qualitative imaging. "People have not even begun to realize just how much more information they're missing in biology because our eyes are unable to see it."

Live imaging

Scott Fraser at Caltech started exploring live imaging of development about 20 years ago. When he and his colleagues used live imaging to follow cell lineages as an animal develops, they found that "the answers we got from this imaging approach, where we're actually watching things as they happen, often were different from the answers we thought we had based on *in situ* patterns or based on fixed tissue," he says. One of the most important discoveries they made in live imaging was that one cell's behavior often depends on its neighboring cells and their behavior. "That's exactly what you can't capture in a culture dish or in a homogenate," Fraser says.



The zebrafish otic vesicle, pictured here at 29 hours (left) and 34 hours (right) post-fertilization, will form the inner ear. Green fluorescence marks cell nuclei, and red fluorescence labels cell membranes. The ear's sensory hair cells do not appear until the more advanced stage (right). The blue spot is the otolith, a small structure attached to hair cells that detects gravity and vibrations. Each image is magnified 40 times. Images: Sean Megason, California Institute of Technology.

After years of perfecting imaging techniques in single embryos, "our work had matured to a point that we thought we could do these sorts of things not one at a time or two at a time, but at a high enough throughput that it made sense to do a more organized effort," Fraser says.

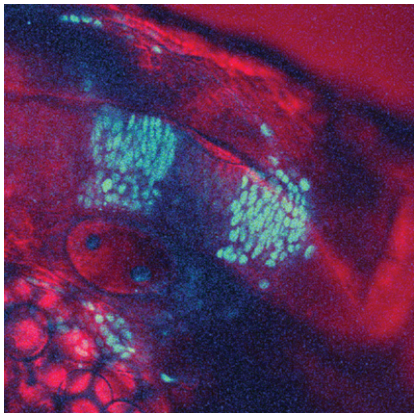
In August 2006, Fraser and Megason, along with fellow Caltech researchers Marianne Bronner-Fraser and Niles Pierce, received a five-year, $18 million grant from the National Institutes of Health's National Human Genome Research Institute. With this money, they've created the Center for In Toto Genomic Analysis of Vertebrate Development at Caltech. The center's goal is to image every developmentally important vertebrate gene. To do this, the Caltech scientists are developing a new type of imaging that they call *in toto* imaging.

"The idea of *in toto* imaging is to be able to image all the cells in the animal at single-cell resolution and then do that over time," Megason says. He and his colleagues are imaging cellular movements and divisions, as well as gene and protein expression patterns, in the developing zebrafish embryo.

Computer analyses of these images reveal expression levels of each gene in each cell. Megason, Fraser, and colleagues are also conducting mutant studies to see how mutations in various genes affect zebrafish development. Eventually, these imaging data will be compiled to create a "digital fish"—a computer model of all of the genes, proteins, cell movements, and cell divisions that transform a zebrafish egg into an embryo. "I want to be able to know which cells are brothers and sisters and which are second cousins, and I want to know which genes correlate with those relationships and with their eventual fates," Fraser says.

The developing zebrafish is ideal for live 3-D imaging, Fraser says, because it's small and transparent as it develops, so it's easy to peer inside with a microscope. To image gene expression quantitatively in each zebrafish cell, the researchers have developed a technology called "flip trapping." They insert a tag consisting of a gene for a yellow fluorescent protein into a zebrafish gene of interest. When the gene's protein is expressed during development, the yellow fluorescence tells them where and at what level it is expressed.

In the presence of a certain enzyme, however, the gene mutates, disturbing the associated protein's function. At the same time, the genetic insert changes its formation, and instead expresses a red fluorescent protein. The resulting red fluorescence indicates the cells with the mutant phenotype, revealing something about the gene's role in development.



This FlipTrap transgenic zebrafish embryo, magnified 40 times, is shown at 36 hours postfertilization. The green indicates expression from the FlipTrap in the nuclei of hindbrain cells. The red is transmitted light. Image: Sean Megason, California Institute of Technology.

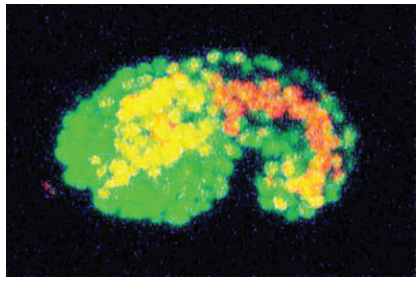
The researchers then use laser-scanning microscopy to image the expression of each fluorescently tagged protein as the animal develops. Their microscopes record time-lapse movies of "the whole three-dimensional structure of an embryo as it's developing," Fraser says. The movies allow them to follow each cell as it develops, turns on and off different genes, and interacts with neighboring cells.

"Our goal is, in the next three years, say, to have more than a thousand genes labeled and mutated," Fraser says. Just five years ago, imaging the expression of 1000 genes in the zebrafish embryo would have seemed like an "impossibly long" project, he says. At that point, imaging the expression of one gene throughout development would have taken at least six months. Now, it can be done in about a week.

In the first year of the project, the researchers are working on "refining the technology to see if we can get another tenfold improvement in every step," Fraser says. Starting next year, they'll go into full production mode. They plan to image the expression of a few hundred genes per year, along with the family tree of the cells that express those genes. (Next spring, Megason is moving to Harvard Medical School in Cambridge, Massachusetts, to set up his own lab in the systems biology department. The project he has started with his Caltech colleagues "will go on, but will now be bicoastal," he says.)

After they've made progress in the zebrafish, the Caltech researchers also plan to image development in the quail. "We wanted a higher vertebrate that develops in a way very similar to humans," Fraser says. They'll eventually apply the same imaging and mutagenesis techniques to the quail that they're now using in the zebrafish. They've already had success "labeling individual cells and collecting data on small groups of cells at a time," so they know they can do the imaging, Fraser says.

A group at the University of Washington (UW) in Seattle is also using live, single-cell imaging to decipher the developmental genes and pathways in the nematode *C. elegans*. They have developed an automated system that traces cell lineage during embryogenesis. *C. elegans* is ideal for this type of project because it has an invariant cell lineage and fate, says Zhirong Bao of UW. Every adult has exactly 959 somatic cells, and the developmental pathways of each of these cells have been mapped out. "If you can trace the lineage of a cell, you not only know where it's from, you also know exactly what it's going to give rise to," Bao says. "So once you map gene expression or phenotype onto a particular cell, you've mapped it to a well-defined developmental context."



Red fluorescence in this Caenorhabditis elegans *embryo, magnified 630 times, marks expression of* pha-4, *a gene encoding a transcription factor found in the digestive tract. Green fluorescence marks cell nuclei. Image: Zhirong Bao, John Murray, and Bob Waterston, University of Washington.*

Because of the worm's small number of cells and well-defined developmental program, Bao is hopeful that *C. elegans* will be the first organism in which scientists will know the expression of "every gene in every cell throughout its life. That's the ultimate goal for this project." Bao and his colleagues, including John Murray and Bob Waterston, also of UW, label each cell by attaching GFP to a core nuclear protein and then use time-lapse imaging to take single-cell resolution pictures of the developing worm. Their image-analysis algorithms then process the images, tracking the cells' movements, divisions, and deaths as the animal develops.

Lineaging algorithms are not perfect—the researchers still have to correct some errors manually—but the programs are constantly being improved, Bao says. His group's software can currently track cells through eight rounds of cell division with more than 99 percent accuracy and through nine rounds with more than 90 percent accuracy. The researchers are working on getting the program to track the 10th and final round of cell division in *C. elegans*. Cell tracking becomes more difficult as more cells appear and become densely packed, Bao notes. "It's still work in development. We hope that, in the end, we can do close to 100 percent throughout embryogenesis."

The UW researchers are now using their system to detect gene expression levels in single cells by attaching a second fluorescent tag to genes they're interested in. So far, they've measured gene expression throughout development for about a dozen genes. Waterston, Bao's advisor, is heading a consortium that plans to map gene expression at single-cell resolution for all 900 transcription factors in the *C. elegans* genome in the next five years.

The system should also allow researchers to observe how mutations in certain genes affect cell lineage development. "We know the position and movement of every cell at every minute," Bao says, so "we have the power to look at phenotype at single-cell resolution."

The *Drosophila* network

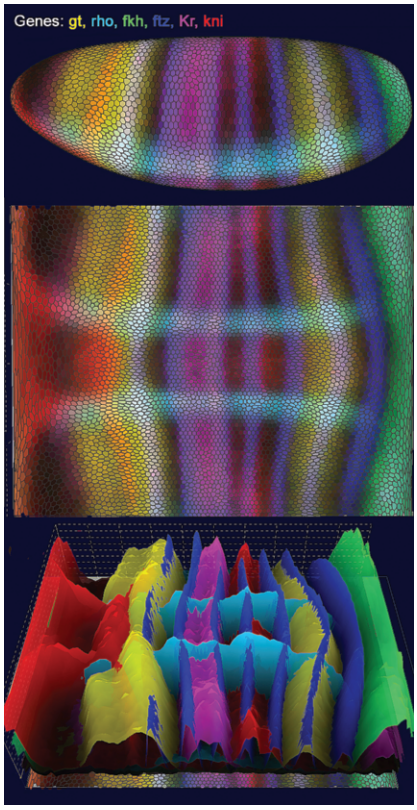
At the Berkeley *Drosophila* Transcription Network Project, researchers at the University of California–Berkeley, the Lawrence Berkeley National Laboratory, and the University of California–Davis are using similar imaging techniques to study development in the fruit fly. "[The] overall goal of the whole project is to model animal transcription networks," says Biggin, one of the principal investigators on the project. As with the Caltech and UW work, the researchers' goal is to measure quantitative gene expression at a cellular resolution.

Unlike the the scientists working on the zebrafish and *C. elegans* projects, however, the Berkeley researchers are doing most of their work in fixed tissue rather than in live embryos. The size and opaqueness of the fly embryo make it impossible to image clearly through the entire organism, Biggin explains. "With live-cell [imaging], we can't see all of the *Drosophila* embryo. With fixed-cell [methods], we can go in with one embryo and capture all the cells."

For now, the researchers are focusing their efforts on the first 100 minutes of gene transcription in a new embryo. They're measuring single-cell expression of about 40 primary transcriptional regulators and 1000 of their target genes, which control pattern formation in the early *Drosophila melanogaster* embryo.

The researchers fix embryos at different developmental stages and then label each cell by staining nuclear DNA with a green fluorescent marker. They use LSM to image a stack of about 100 2-D slices through the embryo, reassembling them to create a 3-D picture. The image software recognizes that each clump of green pixels corresponds to the nucleus of one cell, and the program makes a note of the position of each cell in the embryo. Because the project uses mainly fixed instead of live embryos, the researchers have also conducted experiments in live animals to confirm their findings. "The fixed material is preserving very well things seen in living material," Biggin reports.

The scientists also attach blue or red fluorescent tags to messenger RNA molecules or proteins to image cellular expression levels of two genes in each embryo they image. "We get a quantitation for each cell of the relative level of gene expression," Biggin says. Each cell's location and gene expression levels, as well as information regarding the embryo's orientation, developmental stage, and morphological features, are compiled into text files the researchers call PointClouds. "Because of this image analysis, we can extract [all the image information] into basically a spreadsheet file that you can run on your laptop," says David Knowles, another principal investigator from Lawrence Berkeley National Laboratory. "You can watch all these genes move in time and space."



The messenger RNA expression of 5 genes in the early Drosophila *embryo is shown using a visualization tool called PointCloudExplore, which was developed by researchers at the Berkeley* Drosophila *Transcription Network Project. The genes displayed have been selected from 95 available in the current* Drosophila *expression atlas. The top image is a three-dimensional model of gene expression; the middle image shows a two-dimensional cylindrical projection of that expression; and the bottom image is a map of the cylindrical projection, where height represents the level of gene expression in each cell. Image: Berkeley* Drosophila *Transcription Network Project.*

Eventually, the Berkeley researchers hope to create a single computer file of a developing fly embryo, which scientists can use to watch and analyze expression patterns of hundreds of genes as they change simultaneously. "We're building an atlas of all the gene expression patterns that can be mapped onto one coordinate frame," Knowles says, which he and his colleagues call a "virtual embryo." The researchers have so far compiled a digital atlas of expression of about 100 genes at various stages of early embryo development, which can all be seen in the same virtual embryo.

Last December, the Berkeley researchers published their first results from the project in *Genome Biology*. By imaging the expression of 22 genes in 1282 fly embryos, they revealed several features of morphology and gene expression that hadn't been seen before in the *Drosophila* embryo. For example, during early development, "stripes" of gene expression move across the embryo. Scientists thought that some cells turned off a particular gene's expression while their neighboring cells turned it on, giving the

Visit these Web sites for more information:

Center of Excellence in Genomic Science, California Institute of Technology: *http://cegs.caltech.edu*

Waterston lab image collection: *http://waterston.gs.washington.edu*

Berkeley *Drosophila* Transcription Network Project: *http://bdtnp.lbl.gov*

appearance of expression movement while the cells themselves remained stationary. However, using their new imaging data, the Berkeley researchers discovered that cells do move during this stage—something that had never been seen before.

Their 3-D models revealed several other novel features of *Drosophila* embryo development. They found that genes thought only to regulate development along the anterior–posterior axis also regulate the dorsal–ventral axis. And they found that the early embryo has an intricate cellular patterning that correlates well with later morphological features of the fly. "We've been able to find that there's this rich, complex pattern of cell morphology in the embryo which had gone unnoticed for literally 100 years of analysis," Biggin says.

In the future, the researchers may apply their methods to later stages of *Drosophila* development, but some improvements in optical imaging technology will be required to image something as large as a larva, Biggin says.

Future improvements and data sharing

While the first results from these projects are trickling out, the researchers are working to improve their technologies so they can image faster in the future. "The biggest challenge for the long term for everybody is just the steady improvement of imaging techniques," Biggin says. "The better the images, the easier the image analysis." Improvements in microscopy will also likely allow scientists to image larger, thicker embryos, he says.

Improving image-analysis algorithms is a big concern for scaling up the Caltech project, says Megason. "We're trying to push on the software so it's better able to recognize objects within images," he says. "It's never going to be perfect, but we want to get it better and better, so it takes less human intervention."

The information that these projects reveal about animal development will be useful to a vast number of researchers, much as genome sequences have been. Most groups probably won't do this level of image analysis themselves, but they will "be able to enjoy the fruits of it," Fraser says. The researchers hope to make it easy for other scientists to access their data by centralizing it in large, freely available databases, much as GenBank has done with genome sequences.

The Berkeley researchers have already placed the virtual *Drosophila* embryo on their Web site, so that other scientists can download their data and use it in their own research. "You can click which time point you want to look at and which gene you want to look at, and you can compare them," Biggin says. "And you do this in 3-D." The scientists at Berkeley have also developed ways for researchers to visualize the data on 2-D plots. "We're not very good at looking at quantitative things in 3-D," Biggin says, "so we have to find abstract ways to look at data so people can see the relationships."

The Caltech researchers are also placing their primary data online for anyone to access, which should allow scientists to use the same raw data for many different types of projects, Fraser says. "I'm going to use it to look at the heart or to look at the spinal cord. And somebody else is going to care about [the] pancreas or how part of the forebrain develops." They're also archiving every mutant strain of zebrafish that they create, so that other researchers can use the same animals for their own studies. "Our strong point is our ability to generate really good data," Megason says. "What you do with those data, I think, is still an open question."

No matter what researchers choose to do with these new data, they can expect to learn a lot in the next few years, Biggin says. Because 3-D, single-cell analyses of development have just begun, "everything we measure is going to be new," he says. "Essentially every question we ask, we make some kind of discovery."

Melissa Lee Phillips (e-mail: mlp@nasw.org) is a freelance science writer in New York City.

doi:10.1641/B570803
Include this information when citing this material.