Cell
Leading Edge
Review

Imaging in Systems Biology

Sean G. Megason[1,*] and Scott E. Fraser[1,*]
[1]Beckman Institute and Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA
*Correspondence: megason@hms.harvard.edu (S.G.M.), sefraser@caltech.edu (S.E.F.)
DOI 10.1016/j.cell.2007.08.031

Most systems biology approaches involve determining the structure of biological circuits using genomewide "-omic" analyses. Yet imaging offers the unique advantage of watching biological circuits function over time at single-cell resolution in the intact animal. Here, we discuss the power of integrating imaging tools with more conventional -omic approaches to analyze the biological circuits of microorganisms, plants, and animals.

Although we have nearly complete genome sequences for most model systems, we are very far from understanding how this genomic code is executed during development to build an organism. Implicit in the genome sequence of an organism is a complete set of instructions for constructing that organism (aside from epigenetic factors). The problem is that we as scientists have no way of directly deciphering the code. For example, we cannot predict a protein's function based on its sequence, we cannot predict when and where a protein will be expressed based on a gene's noncoding sequence, and we cannot predict the effect of removing a gene's function on the system as a whole. Currently, the only code that we can easily decipher is the parts list. Thus, the big challenge ahead is to answer the related questions with respect to the underlying genetic regulatory network: How do these parts interact as a system, and how does this system function to create an organism? We call this exciting area of research in the postgenomic era "systems biology."

Because systems biology is still young, its boundaries remain fluid and its practitioners use a number of different tools from a variety of different perspectives. The birth of the field was made possible by genomics and other high-throughput approaches such as proteomics and microarrays. Not surprisingly, these -omic approaches are still important tools of systems biology and have proven invaluable for identifying and characterizing the components of biological systems on a comprehensive scale. Recently, these approaches have been refined to identify possible interactions among components, such as protein-protein interactions and protein-DNA *cis*-regulatory interactions. Yeast two-hybrid, biochemical pull downs, protein microarrays, and chromatin immunoprecipitation have been used to characterize the network of interactions on a large scale in several organisms (Cusick et al., 2005; Harbison et al., 2004). These approaches are providing a rough draft of the components and interactions that comprise biological networks on an unprecedented scale. However, even the best rough drafts are relatively coarse, error prone, and uninformed by biological dynamics. Because of this, current high-throughput data provides only a starting point from which a number of hypothesis-driven "wet" experiments must be done to determine the function of a system. Recent work on the elaboration and testing of gene regulatory networks (Levine and Davidson, 2005) offer ample evidence of the hard work needed to validate direct *cis*-regulatory interactions among even a small set of target genes involved in a circumscribed phase of development.

We argue here that imaging can play a vital role in systems biology, offering a path from rough static models to more refined, quantitative dynamic models. In vivo imaging can capture quantitative data at single-cell resolution and do so noninvasively as the biological circuit functions, offering insights that cannot be matched using in vitro approaches. With the emergence of automated instrumentation and advanced analysis tools, such intravital imaging has become practical for both hypothesis-driven research and high-throughput discovery science. In hypothesis-driven research, for example, multicolor imaging can be used to monitor the essential nodes of a biological circuit in real time. In high-throughput research, standardized reporters and imaging conditions permit intravital imaging to capture data approaching the same large scale as today's -omic approaches but with improved temporal and spatial resolution. We begin this review by comparing the relative roles for -omics and imaging in systems biology, then we discuss the unique requirements for applying imaging to systems biology, and finally we conclude by offering examples of imaging in systems biology.

The Next Generation of Systems Biology

Conventional -omic approaches excel at the first goal of systems biology: characterizing the structure of biological networks. This requires the identification of all the components of the network as well as the identification of all the interactions between them. Systems level analysis works best when all the components and interactions are defined, as the lack of only a few can cause a model to fail. For this reason, comprehensive approaches such as sequencing, microarrays, and interactomic approaches

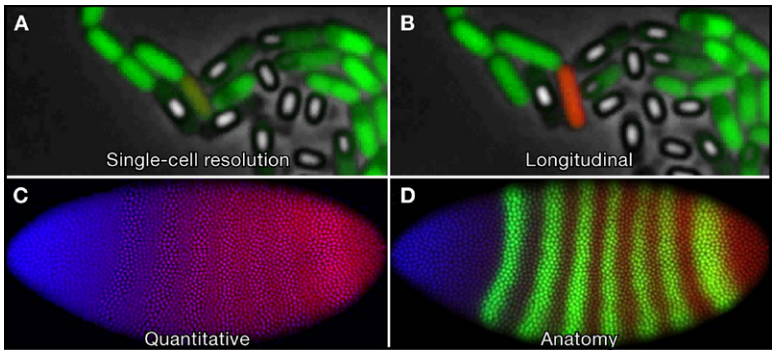

Figure 1. How Biological Circuits Function

There are four important considerations about how biological circuits function that illustrate the advantages of imaging for systems biology.

(A) Biological circuits function at the level of single cells and thus require single-cell resolution for their analysis. Genetically identical cells grown under the same conditions, such as *B. subtillis* in this microcolony, can acquire different phenotypes, such as vegetative growth (green cells) or sporulation (small, light, refractile cells).

(B) Biological circuits function over time and thus require longitudinal analysis such as time-lapse imaging. The same microcolony as in (A) is shown 2 hr later in a time-lapse movie. The circuit for competence (as indicated with a red fluorescent reporter) has been activated in the central cell.

(C) The exact concentration of components in biological circuits is often important for their function and thus requires a technique that is quantitative (preferably at the protein level). Shown is a fly embryo with opposing gradients of Bicoid (blue) and Caudal (red) protein. Bicoid and Caudal are morphogens known to activate circuits differentially as a function of their concentration.

(D) The function of biological circuits can vary across the space of an organism, meaning that their analysis should preserve the anatomical context of the data. Imaging can provide data with high spatial resolution from anatomically intact systems. This is the same embryo as in (C) but stained to reveal expression of Even-skipped, whose seven precise stripes of expression result from the spatial distribution of morphogens. Images in (A) and (B) are reprinted by permission from Macmillan Publishers Limited: Nature (Süel et al. 2006), copyright 2006. Images in (C) and (D) are from FlyEx (Poustelnikova et al., 2004).

are ideal. Genetics can help to identify the components and connections of a network, but redundancy and pleiotropy can mask important links. Once the structure of a network has been elucidated, how it functions as a circuit can be investigated. The advantages of imaging approaches in probing how biological circuits function (Figure 1) are presented below.

Single-Cell Resolution

In studies of biological circuit function, single-cell resolution is a key advantage (Figure 1A). In addition to being the basic building blocks of organisms, cells are the computational unit in biological circuits. Most components of regulatory networks, ranging from signal-transduction components, transcription factors, and *cis*-regulatory elements act within single cells. Although receptor-ligand interactions and secreted signaling proteins can influence distant cells, the interpretation of the signaling is executed within single cells. Thus the network guiding the development of a multicellular organism is best thought of as a network of networks—many intracellular networks linked by a few intercellular interactions. The "output" of many biological circuits is also at the cellular level such as changes in fate, proliferation, apoptosis, or cell shape. Most of these phenomena make little sense at either higher (organismal) or lower (molecular) levels or averaged over a population of cells (a cell cannot be partially dead or partially differentiated). Averaging fails even for populations of cells that at one time appear homogeneous, as genetically identical cells can have significant differences in the expression of components, their connections, and functional states that drastically alter the function of a biological circuit.

Imaging can readily provide subcellular resolution. It is difficult to achieve single-cell resolution with most -omic approaches. Microarrays typically require material from thousands of cells to achieve sufficient signal. Single cells can be analyzed by amplifying the sample, but this might introduce artifacts into the data (Subkhankulova and Livesey, 2006). It is difficult to accurately and reproducibly isolate single cells of a defined stage and position for -omic analysis from a field of cells, and even if successful, the very act of removing the cell can change its context.

Longitudinal Data

Because biological circuits function over time (Figure 1B), techniques are needed to assess the temporal dynamics of the circuit. Yet, most -omic and other biochemical approaches begin by homogenizing the sample, which disrupts the function of the circuit. A time course can be deduced by collecting samples at multiple time points. For example, microarray analysis of the yeast cell cycle has been accomplished by taking aliquots of a culture that was synchronized (through the use of α-factor, size fractionation, or a temperature sensitive cell-cycle mutant [Spellman et al. 1998]). However, such approaches provide population averages that can obscure any fast or variable features of the circuit in the individual cells. Thus, there is simply no way to get adequate longitudinal data using a technique that requires the sample to be destroyed in order to be measured.

Imaging is well suited to collecting longitudinal data because it can be done noninvasively on intact, fully functioning organisms. Time-lapse, fluorescent microscopy can monitor the status of a circuit in each cell of a population repeatedly for hours. Furthermore, imaging can match the timescale of biological computation, which can range from seconds for small molecule signaling (such as Ca^{2+} or cAMP) to hours for protein-based signaling (such as a morphogen-inducing target gene expression).

Quantitation

Rather than being simple on-off switches, biological circuits can show graded responses (Figure 1C). Thus, it is critical to collect quantitative information about the

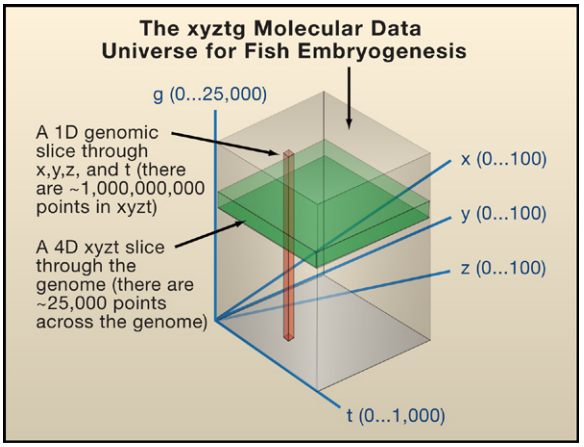

Figure 2. The Universe of Molecular Data
Compared with -omic techniques, imaging can acquire functional genomic data in higher throughput and in a format more amenable to integration. As an example, imagine a data set consisting of the expression pattern of all genes during embryonic development of zebrafish at high enough resolution for use in systems modeling. In rough numbers, such a five-dimensional molecular data universe would contain 25,000 points across the genomic dimension (one for each gene); 1,000 points across the time axis of development (one every 5 min); and 100 points across each of the spatial dimensions of x, y, and z. Time-lapse imaging can capture a 4D slice (xyzt) of this universe for a single gene, whereas -omic methods can only capture a 1D slice at a single point in x, y, z, and t.

circuit's components (such as concentration, localization, and posttranslational modification) to accurately model the function of a circuit. A limitation of many -omic approaches is that they do not accurately measure the quantity of a component. For example, microarrays do not directly provide a measure of quantity but instead provide a relative measure of concentration by comparing the binding of two samples, either through direct competitive hybridization to a single array or by comparing the results from two singly hybridized microarrays. Great effort has gone into developing statistical measures to help interpret such differential microarray measurements (Breitling, 2006). Yet, whether these measurements can be considered quantitative is widely debated in the literature (see Tan et al., 2003; Miklos and Maleszka, 2004; Frantz, 2005; MAQC Consortium, 2006 and references therein). Another limitation of microarrays is that they measure RNA, although proteins do most of the work in biological circuits and it is thus protein concentration that matters most for modeling. Differences in the rates of translation, posttranslational modification, and protein stability limit the correlation between RNA and protein levels ($r \sim 0.4–0.6$). The correlation can vary wildly with the cell type, developmental stage, and category of proteins (Gygi et al., 1999; Greenbaum et al., 2003). Protein and antibody microarrays could address this issue (Kung and Snyder, 2006), but the increased difficulty of working with proteins has made it challenging to develop a robust, specific, and general platform for assaying protein concentration.

Fluorescent imaging has the ability to obtain both qualitative and quantitative data. A fluorescent microscope can serve as a microspectrophotometer, measuring the concentration and distribution of fluorophores inside a cell. Unlike an RNA molecule in a complex mixture hybridizing to a probe on a microarray, there exists a simple linear relationship linking fluorescence and concentration given by the fluorescent dye's extinction coefficient and quantum yield. For example, Wu and Pollard (2005) constructed yeast with yellow fluorescent protein (YFP) fusions to a number of different cytoskeletal proteins expressed from their endogenous genetic loci and showed a very strong linear correlation ($r = 0.99$) between fluorescence and quantitative western blots across a range of expression levels. Thus, the correlation between image intensity and actual number of protein molecules can be much better than that for microarray replicates.

Biological Context

A final consideration in analyzing biological circuits is that anatomy matters (Figure 1D). The structure and function of biological circuits varies from tissue to tissue within an organism, from cell to cell within a tissue, and even between subcellular compartments. Indeed, it is these spatial differences in the function of biological circuits that make different parts of an organism unique. For example, a morphogen diffusing across a field of cells may cause differences in the circuitry depending on the distance from the source. The problem with -omic and other biochemical approaches is that the first step in the procedure is to "homogenize" the sample, destroying the anatomical context of the data, which can blur spatial differences. Imaging has a unique ability to capture data at all these spatial ranges from macro to nano from anatomically intact specimens.

Slicing the Multidimensional Pie

In summary, -omics and imaging can provide insights into many of the same problems, but they do so from very different angles. Typical -omic approaches provide high genomic resolution but low spatial and temporal resolution; in contrast, imaging provides high spatial and temporal resolution but low genomic resolution. To help compare these approaches, it is useful consider the hypothetical "xyztg molecular data universe." Such a five-dimensional data set could describe the expression patterns and subcellular localization patterns for all proteins in the genome (g) at all developmental times (t) and places (xyz) in three dimensions in an organism. Approaches based on -omics and imaging take slices though this data universe in different directions, as depicted in the cartoon in Figure 2. For example, using in toto imaging of a green fluorescent protein (GFP) transgenic organism as described below, it is possible to acquire a 4D (xyzt) slice through the data universe for a single gene. In contrast, microarrays can acquire data across the whole genome but for only a single time and place per array—that is, a 1D genomic slice through many genes but for a single point of x, y, z, and t.

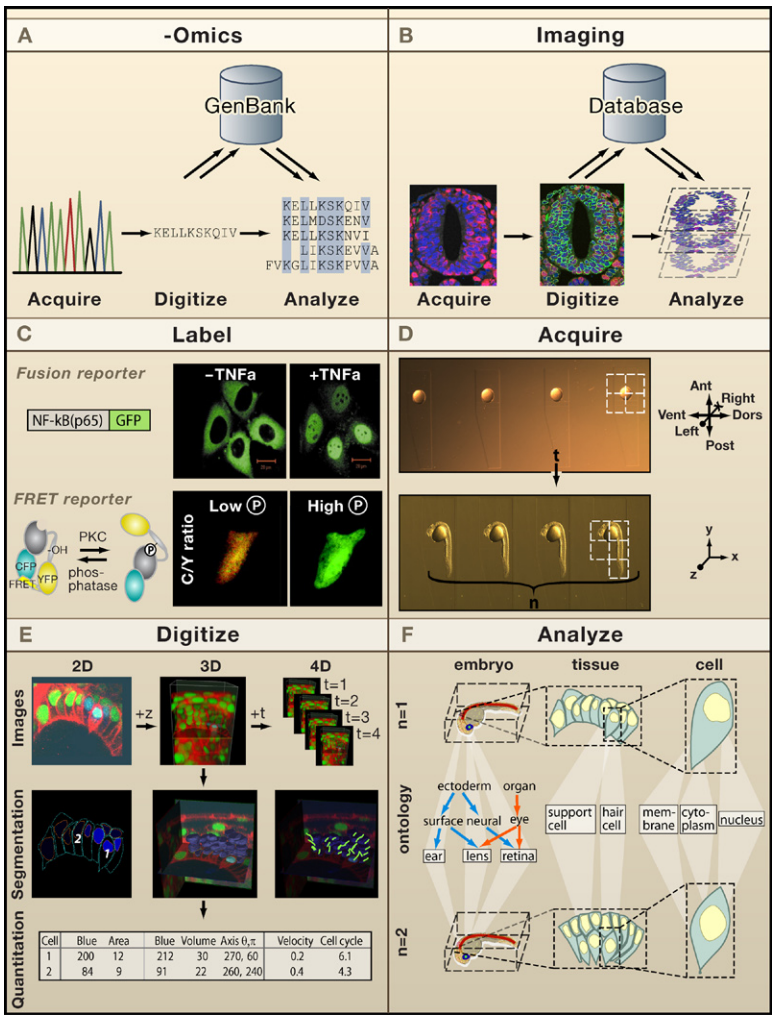

Figure 3. Standardized and Quantitative Digital Imaging

(A) DNA sequences are routinely acquired as fluorescent traces from an automatic sequencer and the bases automatically digitized. These sequences can then be analyzed directly or alternatively uploaded to a central repository such as GenBank. Because sequences in GenBank are standardized, quantitative, digital, and centralized, they can form the basis for systematic, integrative analyses by many subsequent researchers.

(B) This powerful paradigm is typically not followed for other kinds of data such as imaging. To do so requires the data in images to be digitized through the processes of segmentation and quantitation in a standardized fashion. The data can then be stored in central repositories and used for integrative approaches.

(C) Fluorescent fusion reporters can be used to monitor the activity of proteins, which change localization based on functional state, for example the nuclear translocation of the transcription factor NF-κB (Nelson et al., 2002). Reporters based on fluorescent resonance energy transfer (FRET) can report functional changes in proteins (for example phosporylation by PKC).

(D) The next step is to acquire images in a manner that permits systematic analysis, ideally comprehensive across space (xyz) and time (t) and reproducible between samples (n). "Embryo arrays" make this possible for zebrafish by allowing embryos to be positioned in a standardized position and orientation and then imaged at cellular resolution throughout development.

(E) The next step is to extract quantitative, cell-based, digital data from the images. This involves first recognizing cell nuclei and membranes in space as well as cell movements and divisions through time. After segmentation, a wide range of cell-based data can be quantitated.

(F) The final step is analysis. As with DNA, an important issue for both integrating data from multiple samples and creating a centralized database is the use of controlled vocabularies called ontologies. Data from different samples (n) can be registered for comparison if it has been annotated with an ontology. Ontologies can employ different structures such as "is derived from" (blue arrows) or "is part of" (red arrows) and can describe data at a range of anatomical levels from embryos to tissues to cells to organelles. Upper images in (C) are from Nelson et al. (2002) and are reproduced with permission of the Company of Biologists. Lower images in (C) are from Violin et al. (2003), and are reproduced from The Journal of Cell Biology, 2003, 161: 899–909. Copyright 2003 The Rockefeller University Press.

Which transect through the data universe is better for systems biology? Consider throughput, an area in which imaging in typically thought of as low throughput but high content. We believe that the high content can dominate in interrogating the data universe efficiently. Consider what it would take to actually measure the xyztg molecular data universe in zebrafish embryogenesis, where there are roughly 25,000 genes; more than 100 points across each spatial dimension; and 1,000 points across time (roughly cellular spatial resolution and a few minutes of temporal resolution). Because imaging a GFP reporter gene can acquire an entire xyzt slice for a single gene per experiment, it would take 25,000 experiments to capture a complete data set. This is a lot of experiments, to be sure, but not infeasible as the quality, control, and price of microscopes improve. In contrast, -omic techniques could capture the whole genome per experiment but for only a single point in the data cube. To get the same spatial and temporal resolution as 25,000 imaging experiments, it could take as many as 1,000,000,000 (100 x × 100 y × 100 z × 1000 t) -omic experiments. Although this calculation is an exaggeration, as different cells and time points might be pooled for the -omics analysis, it points to the size of the challenge. Even if this were to be optimized, there is the issue of data continuity. It would be very difficult to integrate the 1D genomic data across all the dimensions of x, y, z, and t. In contrast, it would be much easier to superimpose the 4D xyzt data generated for each gene by in toto imaging.

It is unlikely that many researchers will attempt to acquire data on this scale, but similar considerations of resolution, throughput, and continuity apply to more typical, hypothesis-driven experiments. We believe that for many questions in systems biology such as under-

standing the function of a particular biological circuit, the high spatial-temporal resolution from imaging data for a handful of genes is more powerful than the high genomic resolution of -omic approaches for a handful of spatiotemporal points.

How to Image Biological Systems

Although imaging is a widely used technique in biology, its use for systems biology presents a number of challenges. The paradigm shift that started with genomics and continues with systems biology is that data are quantitative, standardized, comprehensive, and (most importantly) digital. Previously, data generated by the scientific community were compartmentalized: they were published in thousands of individual, copyrighted, hard-copy (i.e., analog) journal articles. This made it very difficult to search, reuse, or integrate the huge wealth of data. This was changed by centralized databases that store biological data in a standardized, digital format (Figure 3A). The power of being digital and available is that data can be constantly reused, synthesized, and grown (Negroponte, 1996). Imagine the loss of utility if sequence data only existed in the pages of individual journal articles, that there was no GenBank or BLAST. This power is being extended to some other types of data such as microarray data and in situ expression patterns, but most forms of biological data are still published in an analog form. The availability of standardized, reusable, digital data of many types is essential for the progress of systems biology. Thus, to be useful for systems biology, imaging must be able to capture data in a standardized, quantitative, digital form like that we now take for granted with sequence data (Figure 3B). Capturing data in such a form requires careful forethought of a number of special considerations throughout the pipeline of imaging from labeling to image analysis.

Labeling

One of the most important questions regarding image capture is "what to image"? There are a number of methods for generating contrast in optical imaging, but the most important approach by far is fluorescence. Fluorescent imaging permits a number of different channels to be imaged in one specimen and allows a wide range of very specific structures and even species of molecules to be imaged. Small molecule fluorophores can be used to measure pH and ion concentrations (e.g., Ca^{2+}) and can be used to tag specific proteins by immunofluorescence or with the FlAsH/ReAsH system (Gaietta et al., 2002). The most powerful approach to molecular imaging, however, is through the use of fluorescent proteins such as GFP (Shaner et al., 2005). The main advantage of fluorescent proteins is that they are genetically encoded. This allows them to be used in intact animals for time-lapse imaging, expressed at endogenous levels for quantitative imaging, and genetically engineered to have a wide array of useful properties such as different colors or functional reporters. Fluorescent proteins can be used to mark a variety of things. Transcriptional reporters in which a specific enhancer/promoter is used to drive expression of a fluorescent protein can be used to mark when and where a gene is expressed or to mark specific cell types. Fusion of a fluorescent protein with a protein of interest can be used to characterize its subcellular localization, which can be very important in defining its function, as many proteins change their localization depending on their functional state (Figure 3C). Fluorescent protein fusions allow this change in functional state of a protein to be monitored noninvasively in living cells (Nelson et al., 2002). One concern to keep in mind when working with fluorescent protein transgenics is that its dynamics of degradation may differ from the endogenous protein, especially for transcriptional reporters.

There are other types of functional reporters that can be made with fusion proteins (Giepmans et al., 2006). In fluorescent resonant energy transfer (FRET), a photon can be absorbed by one fluorescent protein and nonradiatively transferred to a nearby fluorophore causing a red shift in emission and often a change in the fluorescent lifetime. FRET only occurs if the fluorescent proteins have overlapping emission and excitation spectra, are close together, and in the proper orientation. FRET reporters, consisting of two fluorescent proteins with a linker containing a recognition sequence for a protease, have been used as in vivo reporters for caspase activity (Xu et al., 1998). FRET reporters can also be made in such a way that a conformational switch in the intervening protein sequence will change the spacing or orientation of the two fluorescent proteins. This approach generally requires detailed structural knowledge of the target protein and some trial and error in the placement of the FPs. Such approaches have been used to generate Ca^{2+} reporters using a calmodulin linker (Miyawaki et al., 1997) and for the detection of phosphorylation by PKC using a linker containing a phosphorylation site and phosphobinding site (Violin et al., 2003). FRET can be used to monitor protein-protein interactions in vivo by fusing one protein with a donor fluorescent protein and the other with an acceptor. This approach has been extended to monitor interactions of 3 proteins using 3-color FRET (Galperin et al., 2004).

It is also possible to construct activity-based reporters using only a single fluorescent protein. Siegel and Isacoff (1997) inserted GFP into Shaker, a voltage-sensitive potassium channel, in such a way that voltage-dependent rearrangements of the channel would alter the fluorescence of the GFP. A final type of fluorescent protein-based reporter uses bimolecular complementation of "split fluorescent proteins," which by themselves are not fluorescent but regain their fluorescence when the two halves are brought into close proximity (Baird et al., 1999; Nagai et al., 2001; Hu et al., 2002).

Image Capture

One of the most significant advantages of imaging for doing systems biology is that it allows data to be captured longitudinally over time through the use of time-lapse imaging. A basic requirement of time-lapse imag-

ing is that the specimen continues to function normally throughout the course of image acquisition. Successfully culturing a specimen is made much more difficult by the requirements of imaging: the specimen must be anesthetized so it does not twitch, immobilized so it does not drift, held in the proper orientation and within the working distance of the objective (which is often short for high numerical aperture objectives), and capable of resisting phototoxicity from the constant bombardment of the photons used for imaging. The requirement of immobilization during culture is fairly easy to meet for some species such as bacteria and yeast, which can be simply imaged on thin agar pads (Elowitz et al., 2002) but is more challenging in other species. Zebrafish embryos can be blocked from twitching with anesthesia and from drifting with dilute agarose (Koster and Fraser, 2004). Chick (Kulesa and Fraser, 2002) and mouse (Jones et al., 2005) embryos can be successfully filmed for 12–36 hr using more involved setups that maintain proper temperature and gas balances while holding the embryo still and near the objective. Photobleaching of the fluorophore and phototoxicity to the specimen are always concerns during time-lapse imaging. These effects can be mitigated by reducing the intensity of the excitation light until images with the minimal acceptable signal-to-noise ratio are produced and through the use of high numerical aperture objectives.

As discussed above, another significant advantage of imaging for systems biology is that it allows biological data to be captured in its full and intact spatial context. For single-celled microorganisms such as bacteria and yeast, a standard wide-field fluorescence microscope is adequate, but for larger organisms it is essential to use a microscope that permits high resolution across the optical axis (depth) in addition to high lateral resolution. In the last decade, two such forms of microscopy, confocal and two-photon, have received widespread use in biomedical research. Both forms of microscopy allow fluorescent imaging of very thin volumes (typically $< 1\ \mu m$) encompassing the focal plane to generate "optical sections." Stacks of optical sections called Z stacks can be collected across a range of focal planes such that entire volumes can be imaged. In confocal microscopy, an adjustable aperture called the "pinhole" is located in the optical path in a conjugate focal plane to the specimen causing most out-of-focus light to be physically blocked. Two-photon microscopy uses a different mechanism to achieve optical sectioning (Williams et al., 1994). Long wavelength (typically infrared) light is used for excitation of the fluorophore by a "two-photon mechanism" in which two photons, each with half the normal excitation energy, are absorbed simultaneously. The chance of two photons being absorbed simultaneously is proportional to the square of the photon flux, which drops off as the inverse square of distance from the focal plane. Thus, the chance of two-photon excitation drops off as the forth power of distance from the focal plane limiting fluorescence to thin optical sections. Lateral resolution is directly proportional to the numerical aperture (a measure of how wide a cone of light an objective captures) of the objective, whereas axial resolution is proportional to the square of the objective's numerical aperture so the use of objectives with high numerical apertures is essential not just for minimizing photodamage but for maximizing resolution. Two-photon imaging can often provide less bleaching and toxicity than confocal microscopy because infrared light interacts less with biological samples and fluorophore excitation is localized to the focal plane (Williams et al., 1994), but this is not a given. Two-photon imaging requires a much greater flux of photons and uses a different region of the spectrum, which may be injurious to some specimens, so the methods must be tested for any system.

Although not yet widespread, several new techniques have shattered the long-accepted resolution limit given by Abbe's Law for far-field optical microscopy (i.e., "normal" microscopes with standard objective lenses). These techniques may make it possible to analyze the location of cellular components with dramatically better resolution than the diffraction limit (Hell, 2007). The possibility that these ultraresolution microcopies will fill in the gap between electron and light microscopy offers exciting prospects for cell and molecular biology.

High-Dimensional Imaging

Combining confocal or two-photon microscopy with time-lapse imaging creates a powerful approach called 4D or xyzt imaging. With 4D imaging it is possible to acquire data with both high spatial context (by imaging intact organisms) and which is longitudinal over time, both of which are important for systems approaches. One of the key ideas introduced by genomics is the value of systematic data, which allows data captured from different organisms to be directly and comprehensively compared, effectively extending the assay across the "n" dimension of sample space. Is it possible to capture images in such a way to cover an entire organism in a standardized, reproducible manner to permit xyzt"n" imaging? For unicellular organisms such as yeast and bacteria, this is straightforward. These organisms can be followed with time-lapse microscopy over several generations, and image series of single cells can be realigned with respect to the cell cycle and cell polarity. Systematic imaging is becoming possible in *C. elegans*. They are small enough that their entire development can be imaged at single-cell resolution and progress is being made to automate the construction of lineages and cell identification (Bao et al., 2006) to allow quantitative comparisons between whole individual animals. Microscopic organisms such as yeast and *C. elegans* can be imaged in a number of orientations and then computationally realigned post facto, but for larger organisms, it is possible to capture the images in a systematic manner from the start. Zebrafish embryos, for example, can be imaged using "embryo arrays" (Figure 3D). Embryo arrays consist of a micromachined plastic template that is used to cast an agarose mold containing an array of

"embryo-shaped" wells. Embryos can then be placed in these wells in a defined position and orientation and time-lapse imaged. The use of a motorized stage on a confocal microscope permits multiple, whole embryos to be imaged across embryonic development in a standardized manner.

Image Analysis

The use of imaging for systems biology requires much more attention to image analysis than needed for many other applications of imaging. In fact, the challenges posed by image analysis are at least as difficult as those faced in image capture. Images are comprised of pixels, but it is generally the objects (e.g., cells) represented by the pixels, not the pixels themselves, that are of interest to a biologist. It is going from a pixel-based representation of data to an object-based representation of data that is the principle challenge in analyzing images. Once objects are recognized in images, it is straightforward for a computer to quantitate their properties such as shape, size, and fluorescence. Recognizing and classifying objects in images is called segmentation (Figure 3E). Although this is a task that the human brain is very good at, computers have a notoriously difficult time.

There are two basic approaches to segmentation: manual and automatic. If there are only a limited number of images to be analyzed, then the objects can be manually identified by the user. If a large number of images must be analyzed, as is often the case when using imaging for systems biology, it is essential to use semiautomated or fully automated segmentation schemes. The difficulty of performing segmentation automatically and the algorithms that should be used varies considerably depending on the cells being segmented. Cells that are of uniform intensity and well separated from each other can be separated by simple thresholding or watershed algorithms (MacAulay and Palcic, 1988). Such algorithms can perform well on cells grown in culture, but they generally perform poorly on cells in tissue because the cells can have uneven intensity and are tightly packed, resulting in some cells being missed and other clumped together in the segmentation. Algorithms that consider cell shape and size to guide the segmentation work better on closely packed cells in tissue (Lin et al., 2003; Dufour et al., 2005).

An important issue is what is being segmented. It is often easier to segment cell nuclei rather than whole cells because nuclei tend to be a more uniform shape and have less apparent overlap with their neighbors, yet segmentation of whole cells is necessary for many experiments, such as quantifying cell shape or total fluorescence. It is important to consider the dimensionality of the segmentation being performed. For simple two-dimensional imaging, there is only one choice (2D) but for xyz, xyt, or xyzt image sets, it is possible to perform the segmentation on 2D, 3D, or even 4D images and to generate 2D, 3D, and 4D segmented objects representing cell volumes and movements over time (Figure 3E). Once cell tracks have been formed to represent the movement of cells, it is often useful to connect the tracks across cell division to trace cell lineage. Computation in biological circuits is often slow, and it is typically necessary to follow cells over multiple cell cycles to observe a circuit in action (Elowitz and Leibler, 2000; Süel et al., 2006).

Currently, there is no universal solution to segmenting cells, cell tracks, and cell lineages from image sets. Algorithms that work well on one set of images often do not work on another set due to differences in cell shape, cell density, the segmentation markers used, signal-to-noise of the images, and resolution. A major goal for software development in the future should be to provide a set of tools capable of extracting cell-based data from many different types of images.

An important issue that is not yet resolved is the need for standards (Figure 3F). Ideally, data generated with imaging for a paper should be preserved for posterity in a centralized, reusable form to serve as a foundation for future work and to facilitate direct comparison between labs. There is a long tradition of submitting new DNA sequences to GenBank as part of publication. For DNA, developing standards such as the GenBank format was fairly straightforward, since DNA itself is digital as it is comprised of a set of four discrete bases. Image-derived data is more difficult to standardize because of its heterogeneity. Although standards have not fully emerged for image-derived data, they should be capable of tying molecular data (e.g., RNA levels and protein levels) to a cell-based framework (e.g., cell type, spatial-temporal coordinates in the organism, and subcellular localization). A major issue here is developing controlled vocabularies called ontologies for describing the data. The Gene Ontology (GO) already provides an accepted vocabulary for describing anatomy below the level of the cell (subcellular localization), and work is progressing on developing anatomical ontologies for cell and tissue types during the development of different animal models (Gene Ontology Consortium, 2006; Baldock and Burger, 2005). A standard for image-derived data should be flexible enough to work across a range of spatial and temporal resolutions, spatial and temporal coverage, and with different markers. Another challenge in creating centralized repositories of image-based data is simply size. The size of image-based data sets can easily be several orders of magnitude larger than sequence-based data sets for typical experiments, creating difficulties for storage in databases and for transmission over the internet. In order to compare microscopy data captured in one laboratory with data from another, it is essential that protocols and reference standards are developed that measure the performance of all the components of the imaging system (Zucker and Price, 2001). Additionally, standards should allow for annotating all the essential technical aspects of the experiment such as culture conditions, objectives, filters, and lasers, similar to what has been done with the Stanford Microarray Database (Demeter et al., 2007).

At the practical level, image analysis for systems biology comes down to software, so what is the best software package? Ideally, a software package would work on very large image sets, allow multidimensional image sets to be visualized, automatically segment and track cells, quantitate the segmented objects, provide methods for using the data for modeling, and provide methods for annotating the data in a standard format. Unfortunately, no such widely useful software exists yet. There are a number of software packages that excel in some of these areas but not others. On the commercial side, the packages of note include Imaris (Bitplane), Amira (Mercury Computer Systems), Volocity (Improvision), and MetaMorph (Molecular Devices). These packages are all for general purpose analysis of biological images and provide tools for multidimensional visualization and some segmentation but are not specialized for systems biology. Academic software packages under active development include ImageJ, 3D-DIAS, CellProfiler, STARRYNITE, Cell-ID, and GoFigure. ImageJ is a Java-based general purpose image analysis program based on the classic NIH Image (Rasband, 2006). 3D-DIAS is designed to automatically segment cells in 3D image stacks of differential interference contrast (DIC) images (Wessels et al., 2006). CellProfiler is designed for high-content screening and systems analysis of cells grown in culture and performs 2D segmentation and cell-based quantitation (Carpenter et al., 2006). STARRYNITE is specifically designed to automatically segment and track cells in the *C. elegans* embryo (Bao et al., 2006). Cell-ID was developed for automatic segmentation and quantitation of fluorescence protein levels in yeast (Gordon et al. 2007). With the above goals in mind, we are currently developing the software package GoFigure. It is designed to automatically segment and track cells in xyzt image sets from a variety of embryos and to extract cell-based quantitative data, which can then be visualized using a number of different linked views.

In Toto Imaging

The logical conclusion of combining the above considerations and techniques is something we term "in toto imaging" (Megason and Fraser, 2003). The goal of in toto imaging is to image and track every single cell in a developing tissue and to digitize all this data in a standardized, quantitative, cell-based manner. In toto imaging can be used for two important processes in systems biology. First, it can be used to construct 4D cell-based anatomical models of all the cell movements, divisions, and shape changes that create parts of an organism. Anatomical models are useful as an armature for constructing models of biological circuits that take into account cell geometry and position. These anatomical models can be thought of as 4D coloring books. Second, data scanned in from transgenics expressing fluorescent proteins can then add color to the cell-based armatures.

Although in toto imaging requires pushing on several areas of technology to become routine, it is becoming possible for small embryos and isolated tissues. We have principally focused on zebrafish because of its suitability for imaging and genetics. For tracking cells, two fluorescent protein markers of different colors are used, one to mark all nuclei through the use of a histone-fluorescent fusion protein and another to mark all membranes with a juxtamembrane-targeted fluorescent protein (Figure 3E). By subtracting the membrane channel from the nuclear channel, the nuclei appear more distinct and thus easier to automatically segment, even in very closely packed tissues. The histone-fluorescent protein marker also permits cells to be tracked over cell division because it stays associated with the chromatin throughout mitosis. Additionally, this dual color approach allows the shape of the entire cell and its subcellular compartments (membrane, cytoplasmic, and nuclear) to be defined. Imaging conditions have been worked out using confocal or 2-photon microscopy to capture 100-μm-deep Z stacks every 2 min for 36 hr, which is sufficient to image every cell movement and division that creates a number of different tissues and organs in the zebrafish embryo. GoFigure is used to process these large (hundred gigabyte) xyzt image sets to construct full cell lineages and extract quantitative, cell-based data. A third color such as an endogenously expressed fluorescent fusion protein can be used to quantitate changes in network components of circuits controlling developmental processes at cellular resolution.

Examples of Imaging in Systems Biology

Below are a few examples to illustrate the points discussed above. This is not meant to be a comprehensive review of these areas but rather to demonstrate the power, versatility, and promise of imaging for the analysis of biological networks.

Bacteria/Yeast

To date the best examples of imaging biological circuits come from single-cell organisms because of their relative ease for both genetic engineering and imaging. Elowitz and colleagues have used imaging to examine the dynamics of both natural and synthetic circuits in bacteria (Elowitz and Leibler, 2000; Elowitz et al., 2002; Rosenfeld et al., 2005; Süel et al., 2006). During times of starvation in *B. subtilis*, most cells make an irreversible choice to form spores, but a small percentage of cells transiently enter into a state of competence in which they can readily take up exogenous DNA. Extensive genetic and biochemical experiments have revealed the major components and molecular interactions that form the circuit regulating sporulation and competence (Stragier and Losick, 1996). However, understanding the dynamic function of a circuit based on its structure is often not straightforward. Furthermore, competence only occurs in a small percentage of cells and in a nonsynchronous manner, making it difficult to study with traditional methods that average over whole populations of cells. Süel et al. (2006) overcame these problems by using time-lapse imaging to watch the dynamics of the competence circuit with single-cell resolution (Figures 1A and 1B). The

core of the circuit controlling competence in *B. subtilis* consists of a master transcription factor, ComK, which is under the control of a direct positive auto-feedback loop and an indirect negative feedback loop through ComS. ComK's output is to activate expression of a number of genes responsible for the competence phenotype including the *comG* operon. The authors used different color fluorescent protein reporters to monitor the activities of the *comK*, *comS*, and *comG* promoters. They found that P_{comK} and P_{comG} expression over time is positively correlated, whereas P_{comS} and P_{comG} expression is negatively correlated. By tracking cells through multiple rounds of cell division, they showed that the chance of a cell entering competence is not dependent on its history or the fate of its sister. These data were interpreted in terms of a mathematical model of an "excitable" circuit in which small fluctuations caused by molecular noise are able to trigger large amplitude excursions of the circuit from its normal steady state. In this case the fast positive feedback loop allows the system to change quickly from one state to another, but the slower negative feedback loop eventually returns the system to the normal vegetative growth state. This model was further supported by altering the circuit to bypass the negative feedback control. Time-lapse imaging showed that these cells exhibit bistability between the normal state and the competent state caused by the positive feedback loop, but now get stuck in the competent state. This work shows how having dynamic, single-cell resolution data for a few genes (limited by the number of colors you can image) can be more useful than having static, population-based data for the whole genome for understanding the function of a circuit. It also shows the importance of being able to track cells longitudinally over multiple rounds of cell division because the timescale for changes in the circuit (i.e., computation) is greater than the length of one cell cycle (12 hr for full induction of competence versus 3.5 hr for normal cell cycle).

Imaging has also been used to study biological circuits in yeast. The chemical reactions that take place inside cells often involve small numbers of molecules and do not occur in a "well-stirred" reaction mixture. These small numbers can lead to stochastic fluctuations or "noise" in biological circuits resulting in phenotypic variation among genetically identical cells. Raser and O'Shea (2004) have characterized the origin of this noise by constructing yeast strains that express cyan fluorescent protein (CFP) from one allele and YFP from the homologous allele. They find that there is both a wide range in the ratio of CFP to YFP within cells (intrinsic noise) as well as the total amount of CFP and YFP between cells (extrinsic noise) and further show that most extrinsic noise is not promoter specific whereas intrinsic noise is (see also Elowitz et al., 2002; Bar-Even et al., 2006; Newman et al., 2006). The thrust of this work is exploring why genetically identical cells within a single population can have phenotypic differences.

Brent and colleagues further characterize the different sources of noise using the yeast mating pheromone pathway (Colman-Lerner et al., 2005). They also use CFP- and YFP-tagged alleles, but instead of imaging cells at only single time points, they use time-lapse imaging to track the change in fluorescence within single cells after induction by mating pheromone. Time-lapse imaging allows them to show that most cell-to-cell variation in response to mating pheromone is caused by initial differences in the capacity of cells to respond rather than noise in the system during its response. Reaching this conclusion required both single-cell resolution and longitudinal data.

To model a network, it is important to not only know all the chemical reactions that define the topology of the network but also to have real data such as protein concentrations and reaction rates to plug into the model. Getting such data using traditional biochemical approaches is difficult because of the complexity of the reactions, the small amount of material often available, and variability between cells. Rosenfeld and colleagues (2005) used time-lapse microscopy to measure the gene regulation function (GRF) that relates the concentration of a transcription factor to the rate of production of downstream products. By imaging the dynamics of a simple circuit consisting of a YFP-repressor fusion protein controlling expression of CFP in single cells, they were able to calculate the GRF. This study demonstrates that time-lapse imaging can be used to calculate "biochemical" parameters such as protein concentration and reaction-rate constants in vivo.

Plants

Moving from single-celled microorganisms to multicellular organisms tremendously increases the difficulty of both imaging and image analysis. Despite these challenges, work in recent years has built upon the powerful and direct observations of plant growth (Selker et al., 1992) and moved us much closer to understanding the cellular and molecular mechanisms regulating plant development. The Meyerowitz lab has developed methods for performing time-lapse, single-cell resolution imaging of the shoot apical meristem (SAM) of *Arabidopsis thaliana* (Reddy et al., 2004). The SAM is a small group of cells at the growing tip of a plant that forms all of the above-ground structures. For each frame of a time-lapse movie, they submerged the SAM underwater, captured the image, and then removed the water, only to repeat this process for the next frame. This laborious method was employed to allow for the use of a high numerical aperture, water immersion objective while still allowing the plant to grow normally. Fortunately, there is no cell migration in plants, which means that tracking the cells from one frame to the next only requires taking into account the pace of cell division and cell expansion. Because of this, frame rates of several hours can be used in plants. This imaging technique was used to map how cell division and cell expansion are patterned across the SAM to generate groups of cells that go on to form a flower (flower primordia).

Heisler et al. (2005) used this imaging technique with GFP reporters to study the role of the auxin pathway in flower specification. Using fluorescent protein fusion reporters, they show that the intracellular distribution of PIN1, a protein involved in auxin transport across cells, is polarized. The pattern of PIN1 polarization across the SAM predicts a pattern of auxin distribution, which is born out using a florescent reporter for auxin activity. The PIN1 polarization pattern is dynamic and points toward newly forming flower primordia.

Jonsson et al. (2006) incorporate these findings with mathematical modeling to develop a model for phyllotaxis. Phyllotaxis refers to the arrangement of leaves or flowers on the shoot of a plant and is arguably one of the longest-standing processes in biology for which a mathematical explanation has been sought. Plants that follow a spiral pattern of phyllotaxis (such as the well-known pattern of florets in a sunflower head) tend to have a characteristic number of spirals (which is part of the Fibonacci sequence), and the spacing between successive primordia follows the golden ratio (~137.5°). To explain the patterning they first used imaging and image analysis to extract the actual shape and neighbor geometry of all the cells in a SAM allowing them to embed a computer simulation of their equations into a geometry that is anatomically correct. Second, they used confocal imaging of a PIN1-GFP fusion reporter to provide a measure of actual protein localization and concentration for all the cells in their model. By embedding the model in a synthetic growing shoot-like topology, it is able to generate a number of common phyllotactic patterns. This approach is being further extended in the Computable Plant project, which aims to use imaging and computer modeling to study plant development (Mjolsness and Meyerowitz, www.computableplant.org).

Animals

High-content image-based screening has been used very successfully on animal cells grown in culture, for example to screen for effects caused by RNAi in *Drosophila* cells (Perrimon and Mathey-Prevot, 2007) or by small molecules in animal cells (Mitchison, 2005). However, this approach has more in common with imaging microorganisms, as discussed above, and has been recently reviewed (Pepperkok and Ellenberg, 2006). We will thus focus on the use of imaging for developing quantitative models of biological processes in intact animals. Unlike plant cells, animal cells do not have a rigid cell wall. As a result, animal cells are able to form a huge variety of different and often geometrically unusual shapes and can change position with respect to their neighbors very quickly. These differences make animal cells much more difficult to both segment and track. The rapid movement of animal cells means that the temporal resolution of time-lapse movies needs to be two orders of magnitude greater for animals than with plants in order to track cells (minutes per frame in animals compared with hours in plants). This dramatically increased temporal resolution has several implications. One is that each image must be captured very quickly. For xyzt imaging (which will be required for most applications) an entire Z stack must be captured within the interval between frames, meaning that each optical section must be captured in a matter of seconds. A second implication is that the specimen will be subjected to a lot of imaging, which can cause problems with photobleaching and phototoxicity. The final implication is that the sheer number of images resulting from an animal experiment may be ~100 times greater than with a plant experiment (which is much greater than with bacteria/yeast), making image analysis much more difficult and automated analysis even more important.

Despite these added challenges, the use of imaging for studying systems biology in animals has tremendous potential because of the relevance of animal models for understanding animal development and human health. To date, most work in animals has been focused on establishing the technology by using animals with small, transparent embryos that can be easily cultured on a microscope stage, such as sea urchin, ascidians, *C. elegans*, *Drosophila*, and zebrafish. Confocal microscopy has been used in sea urchin to quantitate molecular expression by coinjecting eggs with a tissue-specific GFP reporter construct and a freely diffusible red dye (Damle et al., 2006). The red dye is used to correct for the loss of fluorescence with depth, and the GFP reporter fluorescence can be converted into an absolute protein concentration by comparison with embryos injected with known amounts of GFP protein. Ascidians, a primitive chordate, are emerging as a model for systems biology with the completion of their genome sequence and because of the ease of *cis*-regulatory analysis by electroporation of eggs (Shi et al., 2005). Confocal imaging has been used to capture xyzt movies of ascidian embryos expressing three differently colored fluorescent reporter constructs and to track the movement of various cell lineages (Rhee et al., 2005). Progress has been made in creating software (3D Virtual Embryo) and cell-based, digital reconstructions of whole ascidian embryos (Tassy et al., 2006), setting the stage for cell-based, quantitation of in vivo reporters. *C. elegans* is in many respects an ideal animal model for an imaging-based approach because it is transparent throughout life, there is an extensive amount of knowledge and resources available about its development, and importantly it has an invariant cell lineage, which greatly facilitates integrating data from different experiments onto a cellular framework. Lineage tracing has previously been very laborious in *C. elegans*, but recent progress has largely automated the process. Bao et al. (2006) used a histone-GFP fusion to mark all nuclei in *C. elegans*. Confocal imaging was used to capture a Z stack with 1 μm spacing every 1 min throughout embryonic development. A new software package (STARRYNITE) then automatically recognized the nuclei and tracked their lineage up to the 350 cell stage. The larval and pupal stages of *Drosophila* development are very difficult to image, but during the embryonic stage, flies are fairly transparent and suitable for

imaging. The Berkeley *Drosophila* Transcription Network Project is taking advantage of this to conduct a systems analysis of *cis*-regulatory interactions using fluorescent in situs and transgenic reporters (http://bdtnp.lbl.gov/).

Zebrafish embryos are ideal for imaging, in that they are small, easy to culture, transparent, and have very little autofluorescence and scattering (even compared with the embryos described above). As vertebrates, their biological circuits should be much more similar to ours, and they are well suited for genetic approaches. And finally zebrafish develop directly without any larval intermediates, meaning that the structures that form during embryonic development (when imaging is possible) give rise to the actual adult structures. As part of the Center of Excellence in Genomic Science at Caltech, we are currently developing a large collection of fluorescent reporter lines and the imaging and image analysis technology to perform systems-level analyses of vertebrate development (www.digitalfish.org).

Conclusion

Understanding life will require comprehending how the information encoded by a genome is used to create an organism, which can pass this information on to the next generation. It is only now that we have the complete sequence of many genomes that we can begin to approach biology as an information science. We argue that imaging is a particularly valuable tool in this pursuit because of its unique ability to extract information from intact living systems. Although still technically difficult, we believe that the rapid progress in developing new technologies will ensure that imaging plays a vital role in the future of systems biology.

REFERENCES

Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (1999). Circular permutation and receptor insertion within green fluorescent proteins. Proc. Natl. Acad. Sci. USA *96*, 11241–11246.

Baldock, R., and Burger, A. (2005). Anatomical ontologies: names and places in biology. Genome Biol. *6*, 108.

Bao, Z., Murray, J.I., Boyle, T., Ooi, S.L., Sandel, M.J., and Waterston, R.H. (2006). Automated cell lineage tracing in Caenorhabditis elegans. Proc. Natl. Acad. Sci. USA *103*, 2707–2712.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y., and Barkai, N. (2006). Noise in protein expression scales with natural protein abundance. Nat. Genet. *38*, 636–643.

Breitling, R. (2006). Biological microarray interpretation: the rules of engagement. Biochim. Biophys. Acta *1759*, 319–327.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. *7*, R100.

Colman-Lerner, A., Gordon, A., Serra, E., Chin, T., Resnekov, O., Endy, D., Pesce, C.G., and Brent, R. (2005). Regulated cell-to-cell variation in a cell-fate decision system. Nature *437*, 699–706.

Cusick, M.E., Klitgord, N., Vidal, M., and Hill, D.E. (2005). Interactome: gateway into systems biology. Hum. Mol. Genet. *14*, R171–R181.

Damle, S., Hanser, B., Davidson, E.H., and Fraser, S.E. (2006). Confocal quantification of cis-regulatory reporter gene expression in living sea urchin. Dev. Biol. *299*, 543–550.

Demeter, J., Beauheim, C., Gollub, J., Hernandez-Boussard, T., Jin, H., Maier, D., Matese, J.C., Nitzberg, M., Wymore, F., Zachariah, Z.K., et al. (2007). The Stanford Microarray Database: implementation of new analysis tools and open source release of software. Nucleic Acids Res. *35*, D766–D770.

Dufour, A., Shinin, V., Tajbakhsh, S., Guillen-Aghion, N., Olivo-Marin, J.C., and Zimmer, C. (2005). Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces. IEEE Trans. Image Process. *14*, 1396–1410.

Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature *403*, 335–338.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science *297*, 1183–1186.

Frantz, S. (2005). An array of problems. Nat. Rev. Drug Discov. *4*, 362–363.

Gaietta, G., Deerinck, T.J., Adams, S.R., Bouwer, J., Tour, O., Laird, D.W., Sosinsky, G.E., Tsien, R.Y., and Ellisman, M.H. (2002). Multicolor and electron microscopic imaging of connexin trafficking. Science *296*, 503–507.

Galperin, E., Verkhusha, V.V., and Sorkin, A. (2004). Three-chromophore FRET microscopy to analyze multiprotein interactions in living cells. Nat. Methods *1*, 209–217.

Gene Ontology Consortium. (2006). The Gene Ontology (GO) project in 2006. Nucleic Acids Res. *34*, D322–D326.

Giepmans, B.N., Adams, S.R., Ellisman, M.H., and Tsien, R.Y. (2006). The fluorescent toolbox for assessing protein location and function. Science *312*, 217–224.

Gordon, A., Colman-Lerner, A., Chin, T.E., Benjamin, K.R., Yu, R.C., and Brent, R. (2007). Single-cell quantification of molecules and rates using open-source microscope-based cytometry. Nat. Methods *4*, 175–181.

Greenbaum, D., Colangelo, C., Williams, K., and Gerstein, M. (2003). Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol. *4*, 117.

Gygi, S.P., Rochon, Y., Franza, B.R., and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. Mol. Cell. Biol. *19*, 1720–1730.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J., et al. (2004). Transcriptional regulatory code of a eukaryotic genome. Nature *431*, 99–104.

Heisler, M.G., Ohno, C., Das, P., Sieber, P., Reddy, G.V., Long, J.A., and Meyerowitz, E.M. (2005). Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the Arabidopsis inflorescence meristem. Curr. Biol. *15*, 1899–1911.

Hell, S.W. (2007). Far-field optical nanoscopy. Science *316*, 1153–1158.

Hu, C.D., Chinenov, Y., and Kerppola, T.K. (2002). Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol. Cell *9*, 789–798.

Jones, E.A., Baron, M.H., Fraser, S.E., and Dickinson, M.E. (2005). Dynamic in vivo imaging of mammalian hematovascular development using whole embryo culture. Methods Mol. Med. *105*, 381–394.

Jonsson, H., Heisler, M.G., Shapiro, B.E., Meyerowitz, E.M., and Mjolsness, E. (2006). An auxin-driven polarized transport model for phyllotaxis. Proc. Natl. Acad. Sci. USA *103*, 1633–1638.

Koster, R.W., and Fraser, S.E. (2004). Time-lapse microscopy of brain

development. Methods Cell Biol. *76*, 207–235.

Kulesa, P.M., and Fraser, S.E. (2002). Cell dynamics during somite boundary formation revealed by time-lapse analysis. Science *298*, 991–995.

Kung, L.A., and Snyder, M. (2006). Proteome chips for whole-organism assays. Nat. Rev. Mol. Cell Biol. *7*, 617–622.

Levine, M., and Davidson, E.H. (2005). Gene regulatory networks for development. Proc. Natl. Acad. Sci. USA *102*, 4936–4942.

Lin, G., Adiga, U., Olson, K., Guzowski, J.F., Barnes, C.A., and Roysam, B. (2003). A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks. Cytometry A *56*, 23–36.

MacAulay, C., and Palcic, B. (1988). A comparison of some quick and simple threshold selection methods for stained cells. Anal. Quant. Cytol. Histol. *10*, 134–138.

MAQC Consortium. (2006). The MicroArray quality control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat. Biotechnol. *24*, 1151–1161.

Megason, S.G., and Fraser, S.E. (2003). Digitizing life at the level of the cell: high-performance laser-scanning microscopy and image analysis for in toto imaging of development. Mech. Dev. *120*, 1407–1420.

Miklos, G.L., and Maleszka, R. (2004). Microarray reality checks in the context of a complex disease. Nat. Biotechnol. *22*, 615–621.

Mitchison, T.J. (2005). Small-molecule screening and profiling by using automated microscopy. ChemBioChem *6*, 33–39.

Miyawaki, A., Llopis, J., Heim, R., McCaffery, J.M., Adams, J.A., Ikura, M., and Tsien, R.Y. (1997). Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. Nature *388*, 882–887.

Nagai, T., Sawano, A., Park, E.S., and Miyawaki, A. (2001). Circularly permuted green fluorescent proteins engineered to sense Ca2+. Proc. Natl. Acad. Sci. USA *98*, 3197–3202.

Negroponte, N. (1996). Being Digital (New York: Vintage Publishing).

Nelson, G., Paraoan, L., Spiller, D.G., Wilde, G.J., Browne, M.A., Djali, P.K., Unitt, J.F., Sullivan, E., Floettmann, E., and White, M.R. (2002). Multi-parameter analysis of the kinetics of NF-kappaB signalling and transcription in single living cells. J. Cell Sci. *115*, 1137–1148.

Newman, J.R., Ghaemmaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi, J.L., and Weissman, J.S. (2006). Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature *441*, 840–846.

Pepperkok, R., and Ellenberg, J. (2006). High-throughput fluorescence microscopy for systems biology. Nat. Rev. Mol. Cell Biol. *7*, 690–696.

Perrimon, N., and Mathey-Prevot, B. (2007). Applications of high-throughput RNA interference screens to problems in cell and developmental biology. Genetics *175*, 7–16.

Poustelnikova, E., Pisarev, A., Blagov, M., Samsonova, M., and Reinitz, J. (2004). A database for management of gene expression data in situ. Bioinformatics *20*, 2212–2221.

Rasband, W.S. (2006). ImageJ (http://rsb.info.nih.gov/ij/).

Raser, J.M., and O'Shea, E.K. (2004). Control of stochasticity in eukaryotic gene expression. Science *304*, 1811–1814.

Rhee, J.M., Oda-Ishii, I., Passamaneck, Y.J., Hadjantonakis, A.K., and Di Gregorio, A. (2005). Live imaging and morphometric analysis of embryonic development in the ascidian Ciona intestinalis. Genesis *43*, 136–147.

Reddy, G.V., Heisler, M.G., Ehrhardt, D.W., and Meyerowitz, E.M. (2004). Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of Arabidopsis thaliana. Development *131*, 4225–4237.

Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S., and Elowitz, M.B. (2005). Gene Regulation at the Single-Cell Level. Science *307*, 1962–1965.

Selker, J.M., Steucek, G.L., and Green, P.B. (1992). Biophysical mechanisms for morphogenetic progressions at the shoot apex. Dev. Biol. *153*, 29–43.

Shaner, N.C., Steinbach, P.A., and Tsien, R.Y. (2005). A guide to choosing fluorescent proteins. Nat. Methods *2*, 905–909.

Shi, W., Levine, M., and Davidson, B. (2005). Unraveling genomic regulatory networks in the simple chordate, Ciona intestinalis. Genome Res. *15*, 1668–1674.

Siegel, M.S., and Isacoff, E.Y. (1997). A genetically encoded optical probe of membrane voltage. Neuron *19*, 735–741.

Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol. Biol. Cell *9*, 3273–3297.

Stragier, P., and Losick, R. (1996). Molecular genetics of sporulation in Bacillus subtilis. Annu. Rev. Genet. *30*, 297–341.

Subkhankulova, T., and Livesey, F.J. (2006). Comparative evaluation of linear and exponential amplification techniques for expression profiling at the single-cell level. Genome Biol. *7*, R18.

Süel, G.M., Garcia-Ojalvo, J., Liberman, L.M., and Elowitz, M.B. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. Nature *440*, 545–550.

Tan, P.K., Downey, T.J., Spitznagel, E.L., Jr., Xu, P., Fu, D., Dimitrov, D.S., Lempicki, R.A., Raaka, B.M., and Cam, M.C. (2003). Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Res. *31*, 5676–5684.

Tassy, O., Daian, F., Hudson, C., Bertrand, V., and Lemaire, P. (2006). A quantitative approach to the study of cell shapes and interactions during early chordate embryogenesis. Curr. Biol. *16*, 345–358.

Violin, J.D., Zhang, J., Tsien, R.Y., and Newton, A.C. (2003). A genetically encoded fluorescent reporter reveals oscillatory phosphorylation by protein kinase C. J. Cell Biol. *161*, 899–909.

Wessels, D., Kuhl, S., and Soll, D.R. (2006). Application of 2D and 3D DIAS to motion analysis of live cells in transmission and confocal microscopy imaging. Methods Mol. Biol. *346*, 261–279.

Williams, R.M., Piston, D.W., and Webb, W.W. (1994). Two-photon molecular excitation provides intrinsic 3-dimensional resolution for laser-based microscopy and microphotochemistry. FASEB J. *8*, 804–813.

Wu, J.Q., and Pollard, T.D. (2005). Counting cytokinesis proteins globally and locally in fission yeast. Science *310*, 310–314.

Xu, X., Gerard, A.L., Huang, B.C., Anderson, D.C., Payan, D.G., and Luo, Y. (1998). Detection of programmed cell death using fluorescence energy transfer. Nucleic Acids Res. *26*, 2034–2035.

Zucker, R.M., and Price, O.T. (2001). Evaluation of confocal system performance. Cytometry *44*, 273–294.