doi:10.1016/j.mibio.2003.10.001

Molecular Imaging and Biology
Vol. 5, No. 5, 312–325. 2003

1536-1632/03 $–see front matter

REVIEW

# NanoSystems Biology

James R. Heath, PhD[1,2], Michael E. Phelps, PhD[2], Leroy Hood, MD, PhD[3]

[1]Department of Chemistry, California Institute of Technology, Pasadena, CA, USA
[2]Department of Molecular and Medical Pharmacology, The David Geffen School of Medicine at UCLA, Pasadena, CA, USA
[3]Institute for Systems Biology, Seattle, WA, USA

Systems biology is an approach in which the digital information of the genome, acted upon by environmental cues, generates the many molecular signatures of gene and protein expression, as well as other, more phenomenological experimental observations. These data may be integrated together to form a testable hypothesis of how a biological organism functions as a system. The central components of systems biology are genetically programmed networks (circuits) within cells and networks of cells. These components establish the organization and function of individual cells and tissues in response to environmental signals such as cell-to-cell communication within organ systems and whole organisms. Within this context, disease is considered as a genetic or environmental reprogramming of cells to gain or lose specific functions that are characteristics of disease. This paper is a combination of three tutorials with an outlined series of technologies, including microfluidics, nanotechnologies, and molecular imaging methods, and we describe how their development should be driven by the needs of systems biology. We also discuss how these technologies can enable a systems biology approach through a pathway from single cells to mouse models of disease and finally to patients. Within this technology base are approaches to develop, use and test molecules as probes that target proteins, DNA and mRNA to test systems biology models, as well as provide molecular diagnostics and molecular therapeutics within a systems biology framework. 

**Key Words:** Nanotechnology; Microfluidics, Systems biology; Positron emission tomography (PET); Molecular imaging; Integrated microfluidics; Molecular probes.

## Introduction

With the sequencing of the human genome effectively completed a few years ago,[1,2] biology entered into what is now called the post-genomic era, with the implication that knowledge of the human genome would fundamentally change the science and practice of biology. As a consequence, biology and medicine stand at perhaps the major inflection point in their history. The Human Genome Project has catalyzed the emergence of two new approaches to biology and medicine: systems biology and predictive and preventive medicine—and the two, as we shall see, are intimately related. Let us first consider systems biology.

Systems biology analyzes biological systems by measuring the steady-state and dynamic relationships of the elements in the system in response to genetic or environmental perturbations across their developmental or physiological time dimensions.[3] Systems biology is concerned with defining the protein biomodules (e.g., groups of proteins that execute a particular phenotypic function such as galactose and glucose metabolism or protein synthesis) and the protein networks of life (e.g., the skeletal framework of cells and their signal transduction pathways). Systems biology is also concerned with delineating the gene regulatory networks that govern the expression patterns of proteins across developmental or physiological time spans. It is also concerned with delineating the cells effective integration of the protein and gene regulatory networks.

Systems biology must capture global sets of biological information from as many hierarchical levels of information as possible (DNA, RNA, protein, protein interactions, biomodules, protein and gene regulatory networks, cells, organs, individuals, populations, ecologies) and integrate them.[4] This is the start point for the formulation of detailed graphical[3] or mathematical

Address correspondence to: James R. Heath, PhD, California Institute of Technology, Chemistry, MC 127-72, 1200 E. California Blvd., Pasadena, CA 91125. E-mail: heath@caltech.edu

models of biological systems, which are then refined by hypothesis-driven, iterative systems perturbations and data integration. The key is that phenotypic features of the system must be tied directly to the behavior of the protein and gene regulatory networks. Ultimately, these models will explain the systems or emergent properties of the biological system of interest. Once the model is sufficiently accurate and detailed, it will allow biologists to accomplish two tasks never before possible: 1) predict the behavior of the system given any perturbation, and 2) redesign or perturb (e.g., with drugs) the gene regulatory networks to create completely new emergent systems properties. This latter possibility lies at the heart of preventative medicine. Thus, systems biology is hypothesis-driven, global, quantitative, iterative, integrative, and dynamic.

Predictive medicine will emerge for the next 10 to 15 years. It is predicated on two ideas. First, variant forms of human genes predispose to certain diseases. For example, when a woman has a single bad copy of the breast cancer 1 gene, she has a 70% possibility of getting breast cancer by the time she is 60 years of age. In the next 10 to15 years, we will be able to sequence individual human genomes rapidly and economically and, accordingly, provide each individual with a probabilistic future health history. Second, during this same time span, we will be able to use hand-held microfluidics devices to analyze tens of thousands of blood elements (e.g., mRNAs, proteins, metabolites) and distinguish health and disease states specifically. This analysis will detect the consequences of both disease predisposing genes and pathologic environmental signals (e.g., infections) and each of these pathologies will have unique molecular signatures. Preventive medicine will use the tools of systems biology to define the protein and/or gene regulatory networks within which pathologic environmental cues or defective genes operate. It will also provide the methods for circumventing the consequences of these pathogenic signals by re-engineering the behavior of these malfunctioning networks through drugs, proteins, genes, or cellular therapies. In this context, let us consider cancer.

It is now well known that virtually all cancers are genetic in origin, and that any class of cancer (breast cancer, for example) stratifies into a variety of diseases, each with their own molecularly distinct signatures of gene expression and protein translation.[4–7] These signatures arise from different combinations of genetic mutations, but lead to the common clinical presentation of, for example, an unexpected tumor mass in the breast. The sequenced human genome has provided a genetic parts list, which is serving as a powerful database against which disease mRNA or protein signatures can be catalogued. However, it could also be argued that this genetic parts list has had very little impact on either the drug discovery process, or on the clinical practice of medicine to date.

As the fundamental knowledge of the genetic signatures of disease grows, the application of systems approaches to disease will revolutionize medicine. The approach of systems biology is to take as an input the many molecular signatures of gene and protein expression, as well as other, more phenomenological information. Network platform programs where the nodes are mRNAs (or proteins) and the edges are properties that interrelate the nodes (e.g., interactions) will be employed to understand the systems basis of disease. As more types of interrelationships are integrated, the structure of the networks will be refined and clarified. The assembled networks constitute a hypothesis of how the system works at the molecular scale.[4] This hypothesis may be represented graphically in the form of interacting clusters, or modules of proteins, and each module is typically assigned primary responsibility for some biological function, such as respiration or reproduction (Figure 1). This hypothesis is then tested by designing and performing a systematic series of top-down and bottom-up perturbations on the system. For example, a top-down perturbation might be to expose the system to a particular molecule to, for example, inhibit a protein node in a cell circuit. A bottom-up perturbation might be a genetic knock-out to, for example, alter the transcription and translation of proteins that are necessary for a signal transduction pathway in a cell. In practice, hundreds or even thousands of such systematic perturbations can be carried out in parallel. The various molecular and phenomenological signatures are again measured, and the hypothesis is appropriately modified to account for these new results. This modified hypothesis is again tested, and so on, until the hypothesis becomes fully predictive. In principle, with a sufficient number of measurements, a complete molecular description of the system is possible. Such a description contains within it not just the molecular signatures useful for identifying the nature and progression of disease, but also the critical molecular nodes of disease that can serve as drug targets. In fact, the top-down molecular perturbations to the system may be viewed also as a pharmaceutical screening procedure. However, instead of screening molecules against a particular molecular (protein) target, for example, the molecules are screened against a particular system function. In order to gauge the response of the system to the molecular perturbation, it may well be necessary to monitor the responses of tens of thousands of molecular signatures.

A fully realized systems biology approach might appear to be an overly ambitious and unwieldy method for either drug discovery or for disease diagnosis. In fact, within the context of current biotechnologies, it is. For

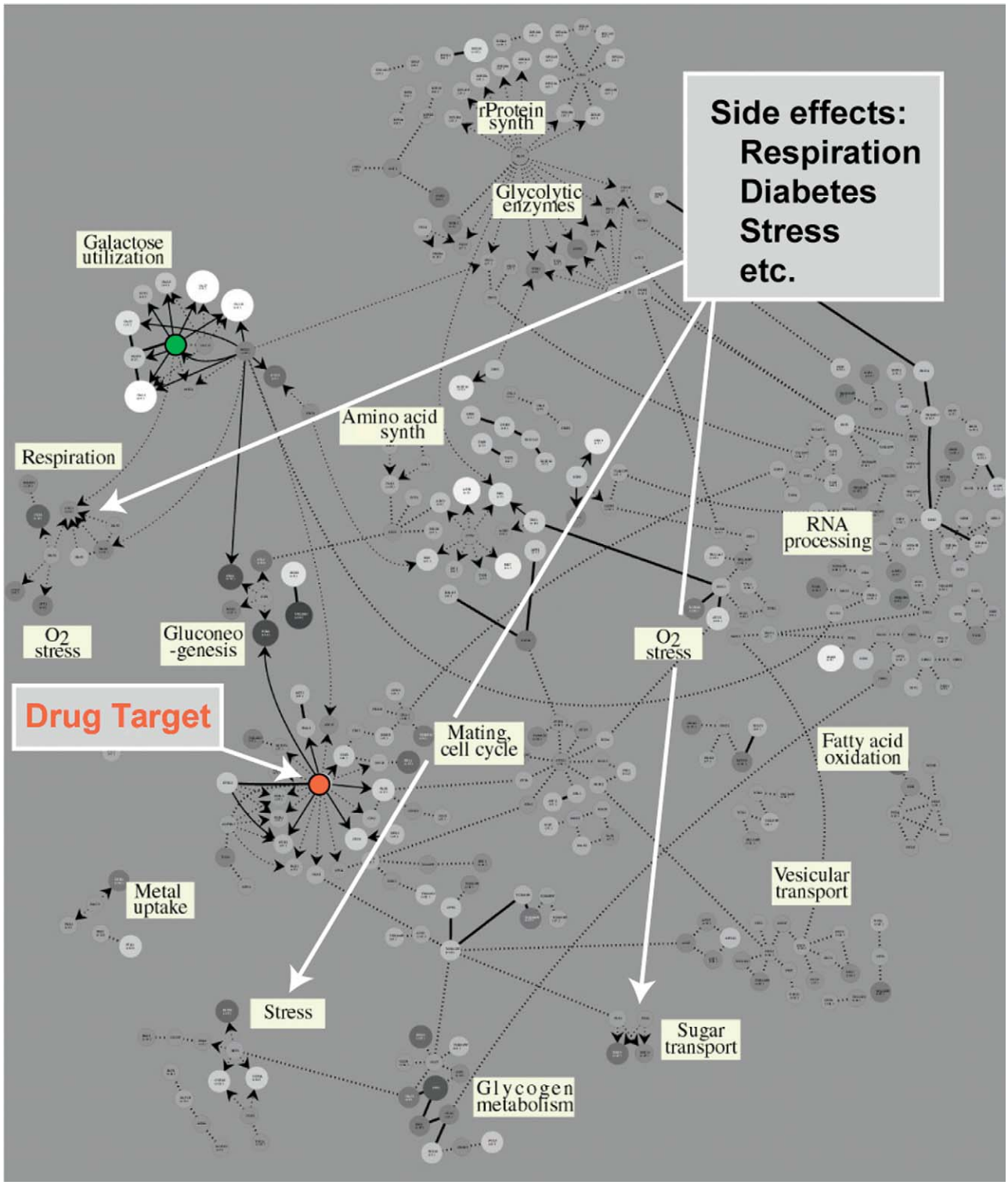


**Figure 1.** A graphical illustration of the protein and gene networks of systems biology in the yeast cell. The various circular shapes represent proteins, and they are grouped according to an experimentally determined set of relationships that arise from a global analysis of protein and mRNA levels and yeast protein/protein and protein/DNA interactions. Each group of proteins represents a protein module that has primary responsibility for a particular biological function. The protein modules are, however, highly interconnected. This map represents a molecular-based hypothesis of how yeast works. This particular hypothesis was tested by a bottom-up perturbation in which the gene that codes for the protein, represented by the green colored circle, was knocked out. All of the proteins shown in this graphic were up or down regulated. Such an analysis is utilized to identify drug targets, as well as the side effects of a molecular perturbation (i.e., a drug) on the system.

example, just the process of cell sorting, polymerase chain reaction (PCR) amplification, DNA arrays, etc., can take a few days, and this is just for a relatively simple analysis of gene expression patterns. A proteomic analysis would require significantly larger cell populations and more time and, of course, many of the proteins (perhaps the majority) would still not be detected. At the very least, a systems biology approach on even the simplest of systems today would constitute a major and expensive research project.

In this paper, we describe an emerging set of molecular and nanotechnologies that are being developed and integrated into a systems biology laboratory. The point is that the needs of systems biology will drive the nature of the tools developed. Also, there will be an attempt to integrate the ability to make multiple interrelated measurements on one nanochip. These are technologies that can potentially analyze many products of gene and protein expression, in real time, and at the level of a single cell or small cell population – i.e., without the need for PCR amplification. When these nanotechnology-based laboratories (nanolabs) are combined with appropriate informatics routines, the resulting output can be a visualization of the molecular signatures of disease and disease progression, resolved at the level of a single cell. The potential for real-time, multivariable analysis implies that the various molecular signals that are recorded can be utilized to predict new experiments in a direct feedback fashion. Coupling these nanolab and informatics platforms with molecular imaging techniques brings about the potential for identifying the critical molecular nodes of disease, and then directly translating that knowledge to the mouse and patient. Our end-goal is to carry out a full systems analysis of a biological organism, and then to utilize that knowledge to regulate, through molecular intervention, critical molecular nodes within a disease cycle, with direct applications to mouse models of disease and patients. Between the extremes, the systems biology models must on one hand be consistent with *in vitro* findings in the expression of DNA and its translation into proteins and at the other end of the spectrum must be consistent with findings in the mouse and most importantly, the patient. This tutorial is an unapologetically visionary picture of what can be done by combining the leading edge advances of these various disciplines, rather than a description of existing applications. However, as we will discuss, early applications of some of these tools to systems biology are occurring and others are on the near horizon.

## Why Single Cells?

A biological system has many layers of diversity, each building upon the previous. These layers traverse from genes to proteins and from there to protein modules, cells, organs, organisms, and so on up to ecosystems. From the aspect of molecular diagnostics and therapeutics, the cell is the fundamental building block. Experimental measurements on particular cellular processes are typically performed on large cell populations. Such experiments lead necessarily to measurements of an inhomogeneous distribution of responses. For example, consider an experiment in which the cell-cycle is to be investigated on a population of cells. All of the cells can be arrested at the initial (G1) phase of the cell cycle, and will remain at that starting point until they are exposed to some specific extracellular signal, such as a mitogen. Let us say that at time = 0 we introduce this signaling molecule, and so initiate the cell cycle process. If we check back in a few hours, we will find that some of the cells have not yet moved beyond their initial stage, while others are variously distributed in the later (S, G2, M) phases of the cycle. In other words, no matter when we stop this experiment—whether after minutes, hours, or days, we will always find an inhomogeneous distribution of cells. It is clearly interesting, from both fundamental and applied perspectives, to understand both the cells that are not dividing as well as those that have accelerated through the cycle. However, unless these processes are resolved at the level of a single cell, it is not possible to elucidate some of the most interesting and critical details of the system. Similar arguments can be made for immunological processes, such as the interactions between bacterial pathogens and cells of the innate immune system, such as macrophages. Once again, we can set a $t = 0$ for this experiment by introducing the pathogens into a colony of macrophages. After a given time period, some macrophages will be highly mobile, others will be consuming and destroying pathogens, and still others will be transferring information about the pathogens to T-cells.

It is of course attractive to be able to carry out at least some of the systems biology perturbations on populations of similar or diverse cells. However, the systems analysis is best done at the single cell level. By carrying out our systems biology experiments on enough single cells, and then by combining this series of homogeneous measurements we will be able to reconstruct the heterogeneous distribution of responses.

## Why Nanotechnology?

The systems biology approach outlined here requires the simultaneous measurement of many signatures of gene and protein expression in real time across the dynamically changing landscape of developmental or degenerative progression or physiological responses. There are three issues that require nanotechnology solutions—large-scale multiparameter analysis, high sensitivity, and temporal resolution. The most sensitive

standard molecular detection method is provided for by the single-photon counting capabilities of fluorescence detection. In fact, single cell gene expression profiling has been demonstrated using optical methods by Singer's group.[8] This work was a tour de force that revealed both the power and the limitations of optical measurements. Using optical multiplexing methods, Singer's group was able to spatially map 11 distinct signatures of gene expression within a single cell. However, there were a few limitations. First, since optical signals photo-bleach, it was not possible to record a time dependence of the expression signals. Second, while this report significantly pushed the multiplexing capabilities of optics, it is still difficult to see how more than 20 or so distinct molecular signatures could be optically identified. Finally, significant engineering of the cells was required in order to generate the fluorescence signals. None of these limitations are intrinsic to nanoelectronic sensors.

For the past few years there have been literature reports of chemical sensors based on single-walled carbon nanotubes (SWNTs)[9,10] or semiconductor nanowires.[11] For these sensors, the chemical event (i.e., the binding of a target molecule to the nanowire surface) changes the conductivity of the nanotube or nanowire, and so the event may be electronically transduced, hereby avoiding the problem of photo-bleaching. It also implies that the number of different molecules that can be sensed is only limited by the number of electrically addressable nanowire sensors that can be fabricated, allowing large-scale multi-parameter analysis, with the caveat that each nanowire has to be encoded with a different molecular probe.

Nanowire sensors can be extremely sensitive. In Figure 2, we illustrate how these devices are thought to work. Although the detailed mechanism is not completely understood, the most accepted model is that the nanowire sensor acts as a sort of transistor. A standard transistor consists of a semiconductor wire, connected at either end to metal electrodes (the source and drain). A third, noncontacting electrode, called the gate, raises or lowers the conductivity of the nanowire (opens or closes the switch) via the application of an electric field. There are alternative ways in which this gating electric field may be applied. For example, changing the chemical environment surrounding the nanowire changes the local chemical potential. This also amounts to changing the local electric field and can gate the nanowire conductivity, thereby providing a basis for chemical sensing. For example, if a single strand DNA is bound to a nanowire as a linker, the binding of the complimentary mRNA will cause a signal to be generated, thereby indicating a specific hybridization has occurred. This sensing mechanism is called chemically gated nanowire sensing.

A completely different mechanism for nanowire sensing involves a nanomechanical signal transduction mechanism. In this case, the nanowires are piezoresistive wires and are fabricated so that the central part of the nanowire is suspended, but with source and drain electrodes connected at either end, as before. The nanowires are again coated with an appropriate molecular linker for the molecule to be detected. When a molecular target (protein or mRNA) binds to the nanowire, the result is that the mechanical coupling (Brownian motion) of the nanowire to the surrounding solution is increased, thereby increasing the piezoresistive electrical noise of the nanowire.[12] Such a sensing mechanism (called mechanically gated) has the advantage that it is completely immune to phenomena such as the charge screening by electrolyte solution of the chemical potential surrounding the nanowire (i.e., the detection mechanism of chemically gated nanowires). The fabrication challenges for mechanically gated nanowire sensors are, however, more severe. One important characteristic of nanowires is that they have a very high surface to volume ratio, and this means that the smaller the nanowire, the larger the relative interaction area between the molecular environment surrounding the nanowire and the electrically conducting core of the nanowire. It turns out that the best chemically gated nanowire sensors are about 5–10 nm in diameter. If they are much smaller, the measured response is too noisy, and if they are much bigger, the sensitivity is reduced.

A large number of nanowire sensors, each chemically encoded with a probe molecule to sense a particular signature of gene or protein expression, can therefore serve as a key part of our systems biology nanolab. Chemically encoding (biopassivating) each nanowire does represent a significant challenge, although there are literature methods based on scanning probe techniques (dip-pen lithography)[13] or electrochemical methods,[14,15] that can be (and have been[16]) applied to this problem.

Another key aspect of the nanolab that is uniquely addressed by employing nanofabrication methods is the concept of real-time systems biology. Certainly there are many molecular processes that can take place within the time scales of $10^{-14}$ to $10^{-3}$ seconds, such as the making and breaking of chemical bonds, the folding of proteins, etc. However, we are more interested in the rates of molecular signaling. Given that molecules are heavily screened from each other within the electrolyte solution of an intracellular environment, the limiting timescale for signaling processes within a cell is molecular diffusion. The time it takes for a molecule to diffuse through a typical 10 micrometer diameter cell is approximately 0.1 seconds, and so we take this as our critical time scale unit. If we wish to retain this level of time-resolution, then we need to not only place our sensor suite within a few micrometers of the cell we are interrogating, but we would also need to have the

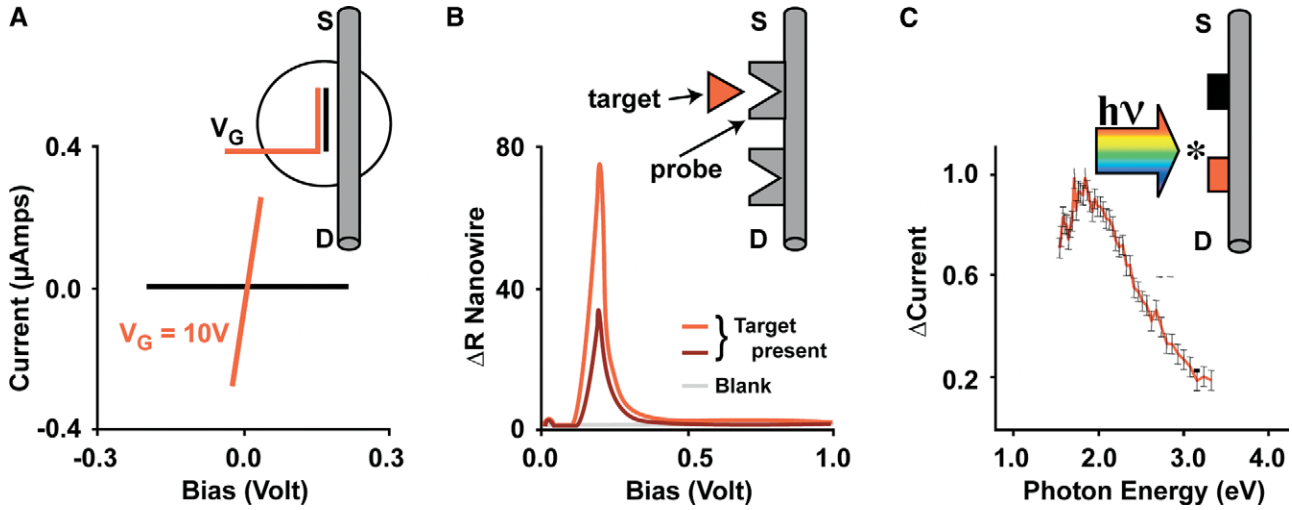


**Figure 2.** Analogous methods for utilizing semiconductor nanowire sensors are shown here, illustrated with experimental data. S and D refer to source and drain contacts, respectively. (A) Normal transistor operation, in which a gate voltage (VG) of 10V turns on electron conductance through the nanowire. (B) Chemical sensing operation, in which a target molecule binds to the nanowire surface, here illustrated as passivated with a probe molecule. This binding event modifies the chemical potential near the surface of the nanowire, again modifying the nanowire conductivity. (C) Optically gated chemical sensing operation, in which a molecule bound to the surface of the nanowire absorbs a photon. The excited state of the molecule exerts a different chemical potential on the nanowire surface, and so again modulates the nanowire conductivity. In this method, the absorption spectrum of molecules bound to the nanowire surface can be electronically transduced.

entire sensor circuit, containing a thousand or more nanowires sensors, to occupy a footprint of only about $10 \times 10$ micrometers squared. This provides a significant challenge. The highest resolution lithographic technique, called electron-beam lithography, cannot pattern semiconducting nanowires at either the diameters or circuit densities that are required. However, a recently developed method, called Superlattice Nanowire Pattern Transfer, or SNAP, does generate the appropriate circuits (Figure 3).[17] Other issues, such as how one goes about electrically addressing each of the nanowires when they are placed together at such high densities,

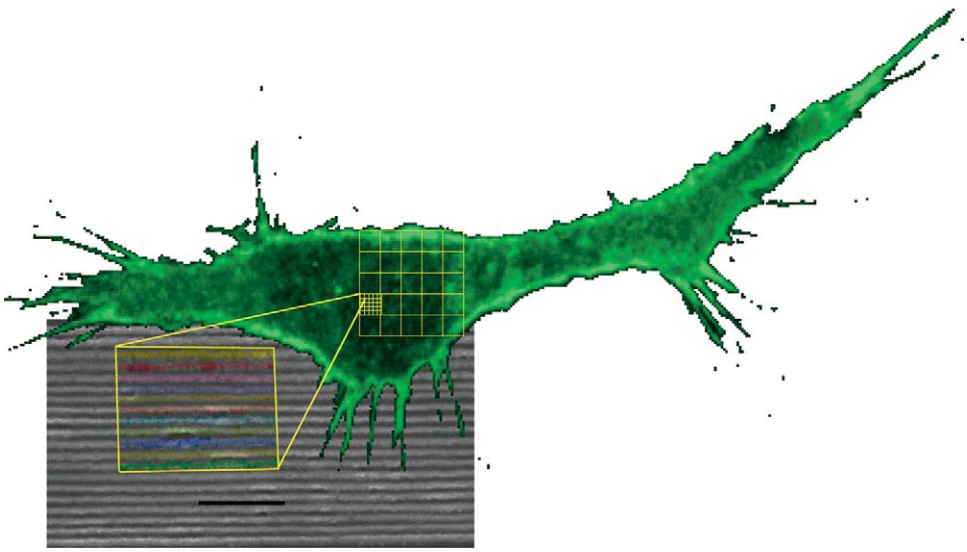

**Figure 3.** A macrophage shown against a suite of 20 nanowire sensors. The yellow square drawn on the macrophage is five micrometers on a side. It is divided into 25 components, and one of those is again so divided. The smallest squares are 200 nm across, and such a square is drawn on the nanowire array. The false colors drawn on the nanowires are meant to indicate that each nanowire is chemically encoded to detect a unique molecular signature. The nanometers are about eight nm in diameter, and separated by about 15 nm. At these patterning densities, approximately 1000 sensors can be constructed onto a footprint the size of a 10 micrometer diameter cell. The macrophage image is courtesy of Dr. Alan Aderem, Institute for Systems Biology, Seattle, Washington.

also provide for significant nanofabrication challenges, but are solvable.

## Microfluidics: Integrating the Biology with the Nanolab

Imagine the following experiments aimed at investigating the cell cycle. First, we stipulate a molecular-based model for the cell cycle that is constructed from existing genomic and proteomic analyses. We test this model by starting with a population of cells that we have arrested at the start of the G1 phase of the cycle. We have transfected these cells with various reporter genes such that as the cells enter each of the four stages of the cycle, they will fluoresce at a different color. We then initiate the cell cycle with some signaling molecule, and we wait some period of time until we can tell, using fluorescence imaging, that there are cells that have progressed into the G1 and the S-phases of the cycle. We then extract several cells from our culture medium, sort them according to their progression through the cycle, and place them into separate nanolab compartments, each containing a suite of electronic (nanowire) sensors. We have separate chemical control over each of the individual nanolab compartments. This enables us to inject small molecules designed to perturb the cells, or introduce nonionic surfactants so that we can investigate cytoplasmic proteins. We can also add ionic surfactants to dissolve the nucleus walls and release the mRNA so that we can measure the signatures of gene expression. At the completion of these experiments, we have carried out a multivariable analysis of approximately 100 cells, with between a few hundred and a few thousand measurements per cell, we have fed those results back into our informatics algorithm, and the initial hypothesis has been modified accordingly. We are now ready to re-test our model by again similarly interrogating and perturbing the cells remaining within our original population.

The above set of experiments would be nearly impossible to execute, even given the various nanotechnology devices and sensors described in the preceding section, were it not possible to automate the entire process. In fact, not only can it be automated, but the entire suite of experiments, from cell culture to cell sorting to cell lysing, etc., can all be incorporated into a single chip-based platform. The key technology for integrating the biology with the nanolab in a highly parallel fashion is a unique and flexible class of microfluidics technologies developed by Steve Quake's group at Caltech[18] and illustrated in Figure 4. They have developed large-scale integrated microfluidic circuitry[19] that is based on multilayer soft-lithography to fabricate micromolded (poly(dimethylsiloxane), or PDMS) elastomer valves, channels, chambers, and pumps. These microfluidics platforms have been demonstrated for cell-sorting applications,[20] (PCR-based) gene-expression profiling at the single cell level,[21] and much more.

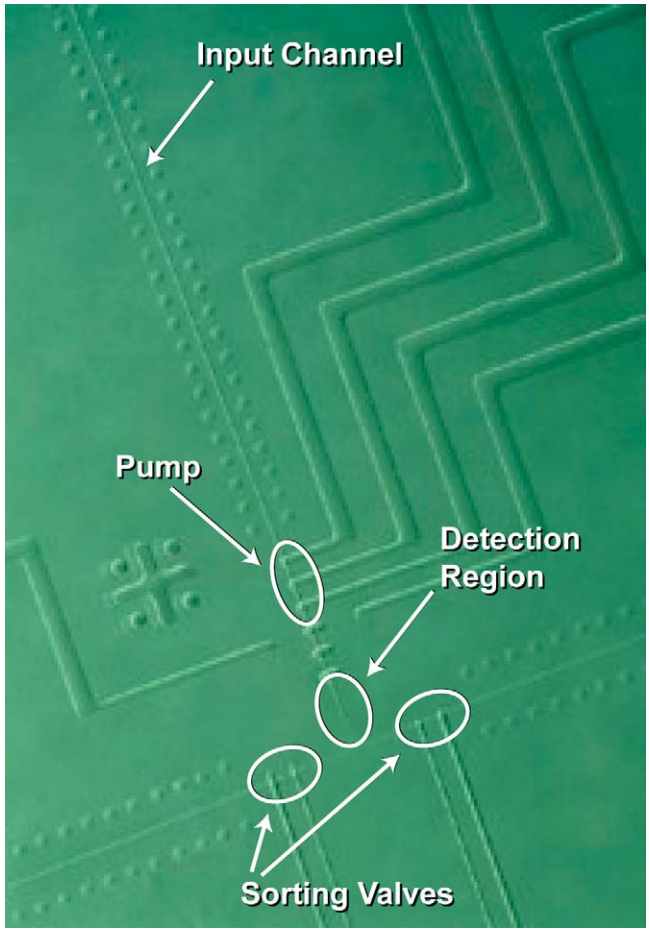


**Figure 4.** One of the components of a microfluidics integrated circuit for plumbing the nanolab: a cell sorter with integrated valves and pumps. This is a two-layer device: the bottom layer is a T-shaped fluidic channel, and the top layer contains pneumatic control lines for pumps and valves, as well as cavities to smooth out oscillations. Photograph courtesy of Dr. Felice Frankel, Massachusetts Institute of Technology.

## Why Molecular Imaging?

Imaging in general implies a technology that can open up a wide aperture to collect signals and generate a picture of structures and/or processes within the field-of-view. A general goal of any scientific endeavor is to collect and process information to form a theory that is illustrated by a picture of the system, such as is presented in Figure 1. The nanolab described above, when coupled with bioinformatics algorithms, will eventually turn into an imaging tool that efficiently collects, processes and displays a molecular-based picture of the biological system being investigated.

By contrast, microscopy and molecular imaging tools begin by presenting the picture. The information content of that picture is nominally based upon the fundamental physical quantity being imaged (i.e., optical fluorescence, surface topography, proton density, radioactive decay, etc.) that are translated into information about biological structures and functions. When coupled with mathematically based data analysis and image reconstruction techniques, imaging methods can be engineered into powerful tools for visualizing key aspects of the biological system under investigation.

A systems biology-derived picture of an organism may be drawn at multiple levels, but has at its origin a molecular basis that connects back to a digital code (the genome) and the molecular machinery that executes that code (the proteome). Thus, if an imaging tool, such as fluorescence microscopy, positron emission tomography imaging (PET), magnetic resonance imaging (MRI), etc., is going to provide a direct connection to a systems biology picture, then that imaging tool, at its most fundamental level, must provide a direct correlation to the molecular basis of the systems biology description. This is, at a minimum, what defines molecular imaging. However, the true benefit of molecular imaging is not that it provides redundant information of the molecular description of the system, but that it can also provide the critical link from the cellular based environment of the nanolab and the *in vivo* environment of the living organism.

The nanolab will eventually evolve into a platform for discovering the molecular probes that can perturb and test the hypothesis of the biological system. The most effective of those probes will be tested within this context as probes for molecular imaging diagnostics and therapeutics from cells and mice to patients using various molecular imaging technologies along this pathway that best serve the purpose.

## A Pathway from the Nanolab on a Chip to Mouse to Patient

As discussed above, the nanolab is an automated experimental environment based upon large-scale integrated microfluidics interfaced to a variety of nanotechnologies. Within this lab there exists a large array of addressable locations and controllable operations in fluid space. Many types of coupled operations and experimental conditions can be laid out within such an interconnected biochemical environment. It is instructive to understand how these various functions can be employed within the context of identifying therapeutic targets and developing both pharmaceutical candidates and molecular imaging probes. The initial goal of the nanolab is to perform large-scale multi-parameter analysis of the molecular basis of systems biology in cells or collections of cells. The second goal is to then reduce this large-scale systems biology problem to a small set of variables that can be used for *in vivo* molecular diagnostics and therapeutics. This reduction in information dimensionality is one of the major challenges of molecular medicine. It is through meeting this goal that we seek to establish the connection between the systems biology experiments and the mouse and patient.

In order to meet this second goal, nanolab capabilities that extend beyond what has already been described are needed. It turns out that many of the engineering problems that are solved by bringing together biology with nanofabrication and microfluidics are general for a host of other applications that are of direct relevance to this goal. For example, additional chips can be devoted to storing molecular libraries, carrying out molecular synthesis, performing protein crystallization, quantifying protein-protein interactions (using nanomechanical devices), etc. In fact, most of these tasks have already been reported in the literature, although many of these applications, such as the carrying out of molecular synthesis within a microfluidic environment, are very poorly developed.

The manner in which these various platforms will be coordinated will certainly vary widely, perhaps limited only by the numbers of practitioners. One possibility, illustrated in Figure 5, might be to initially use the nanolab to carry out a multi-variable analysis of a given system, and to utilize the bio-informatics-generated picture to identify critical protein nodes relevant to a given disease. Those nodes are drug targets, and exposure of the system to a candidate pharmaceutical aimed at those proteins, followed by a multivariable analysis of the response of the system, may provide a compelling pathway for drug testing. It may be necessary to screen many molecules, to chemically modify some of the best initial candidates to optimize the system response, to crystallize the protein in order to assess the nature of the drug candidate-binding, or even to screen the drug candidates against the protein to assess and chemically optimize the interaction forces.

The nanolab will allow for these types of large-scale and rapid multiparameter analyses that are only possible *in vitr*o. However, a good drug candidate or an analog of it, perhaps discovered as described above, may also serve as an excellent molecular imaging probe for diagnostically interrogating the state of the disease in the living organism. Thus, the nanolab becomes a molecular probe (diagnostic and therapeutic) discovery platform as well.

There are various types of molecular and structural imaging techniques that have been and continue to be developed to link *in vitro* imaging to *in vivo* imaging. These include various optical imaging modalities, PET, MRI, single photon emission computed tomography

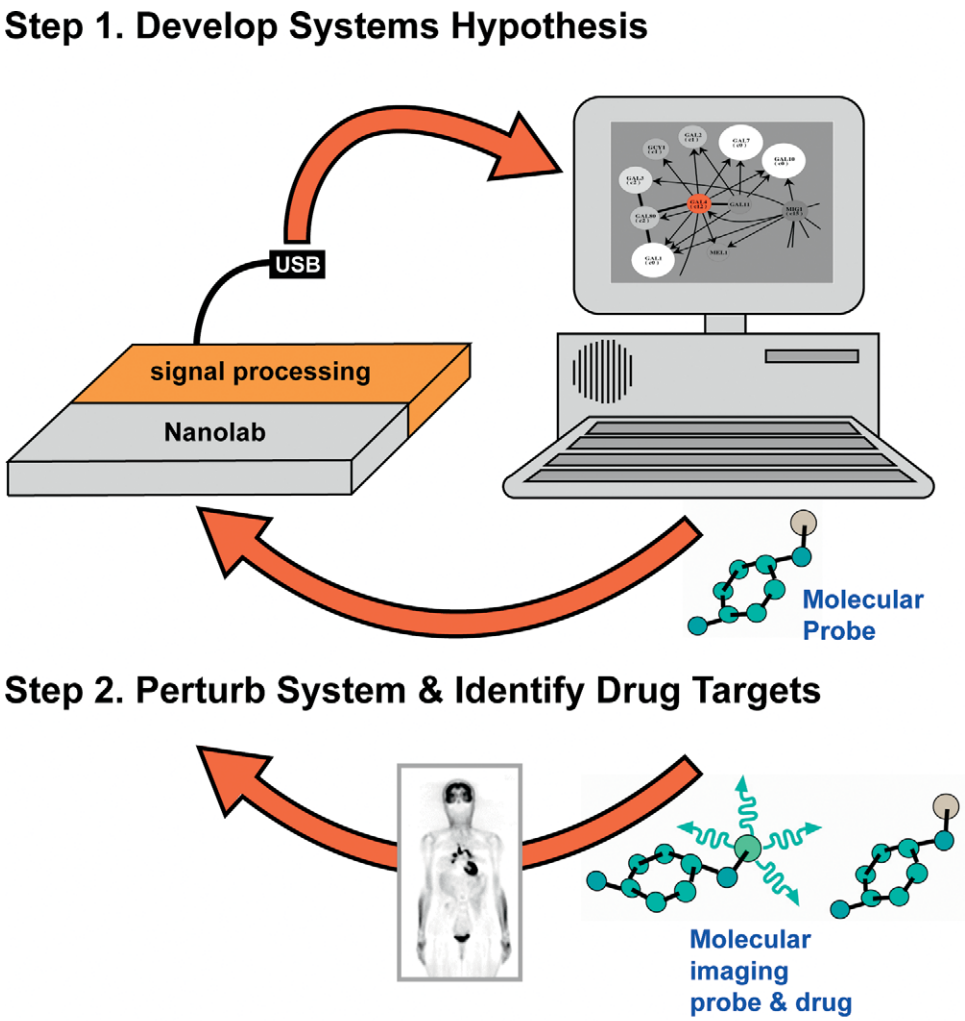


**Figure 5.** Possible methods through which the various tools described in this tutorial might be used within to solve a systems biology problem and to apply that solution to a drug discovery process. In Step 1, the nanolab is utilized to carry out a rapid and informative molecular analysis of a biological system via a global analysis of proteins and mRNA levels in a few cells. This data, coupled with an informatics algorithm, generates a hypothesis of the system. Step 1 also illustrates a means for carrying out an informative molecular-based diagnosis of disease. In Step 2, the hypothesis is tested using top-down perturbations, in the form of molecular probes, and drug targets are identified. In Step 3, the most effective of these molecular probes is turned into an imaging probe (and a drug), and applied toward imaging and treating disease within a living patient.

(SPECT), computed tomography (CT) and ultrasound. Figure 6 illustrates a pathway from *in vitro* studies of cells to *in vivo* examinations of mice to patients that brings together the nanolab, optical imaging, and PET.

Optical imaging provides a molecular imaging link from cells to mice because of its successful and growing use for *in vitro* cell cultures[22] and *in vivo* mouse settings.[23,24] In the mouse, optical imaging is joined with microPET scanners[25,26] for mice to build assays and knowledge from the level of imaging gene expression and DNA replication to signal transduction and metabolism.[27] Figure 7 illustrates the use of optical imaging to examine simple circuit logic of an "AND" logic gate of a protein-protein interaction in a living mouse.[28,29] Figure 8 illustrates a common biological assay with microPET in mice and PET studies in patients to examine alterations in glucose metabolism.

## The Link between Molecular Imaging and Molecular Therapeutics

At present molecular imaging and molecular therapeutics have come together to develop molecular probes that are analogs of each other but have common disease targets.[27] In imaging with PET, the probe is used in extremely low mass amounts (tracers with resulting tissue concentrations of ~ femtomoles/g tissue) to image and measure the concentration and function of the target without mass effects on the target.[26] In the case of the drug, the molecule is given in sufficient amounts to therapeutically alter or block the function of the target. The molecular imaging probe can be used diagnostically or to provide the means to image and measure the pharmacokinetics of the drug in the living mouse and patient, first at tracer levels and then to titrate the drug to selected occupancy on the target. In

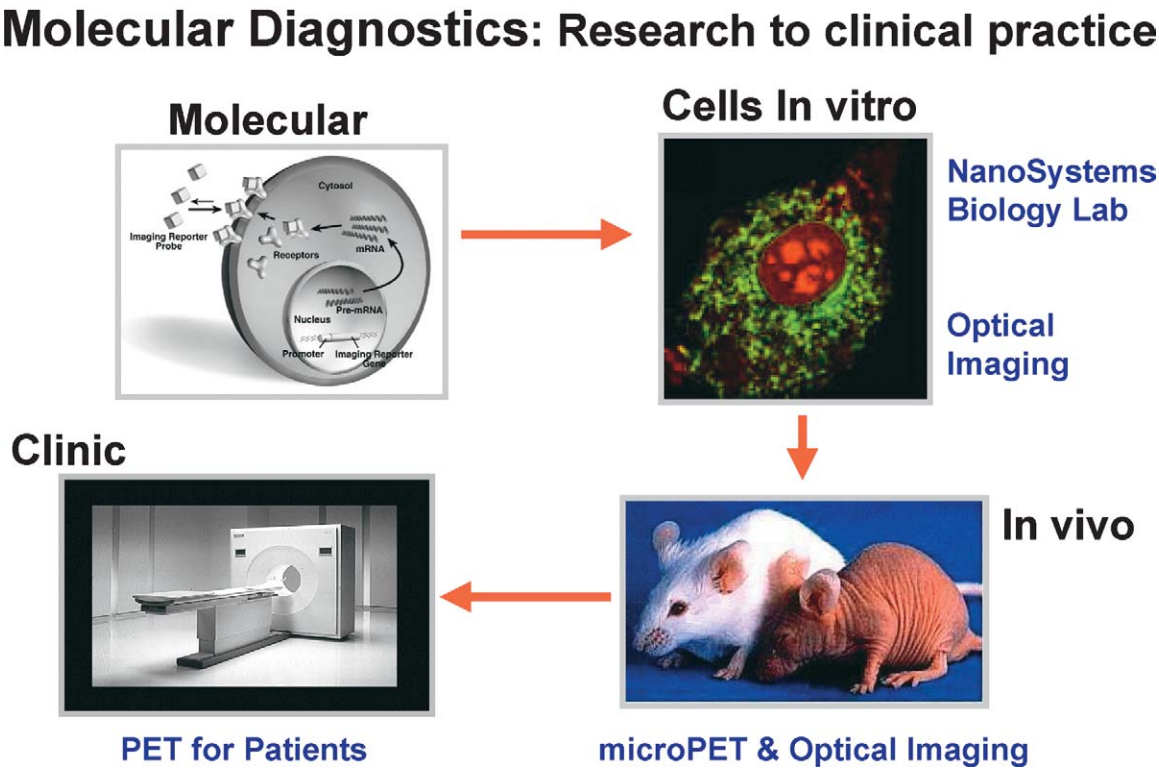


**Figure 6.** Illustration of an experimental pathway from *in vitro* investigations of the molecular basis of cellular processes using the nanolab and optical imaging (image of gene expression in a cell) to the *in vivo* level in the mouse with optical and PET imaging and then to patients with PET.

addition, other PET probes can be used as surrogate markers to examine the drug's effectiveness to modify or terminate a biological process such as metabolism, gene expression, transmitter synthesis, and cell proliferation.[27,30–32]

Figure 9 illustrates the use of PET to monitor the pharmacokinetics of a drug analog. Since PET is a quantitative imaging technique, the concentration of the labeled drug analog can be measured throughout all the organ systems of the body over time (Figure 9).

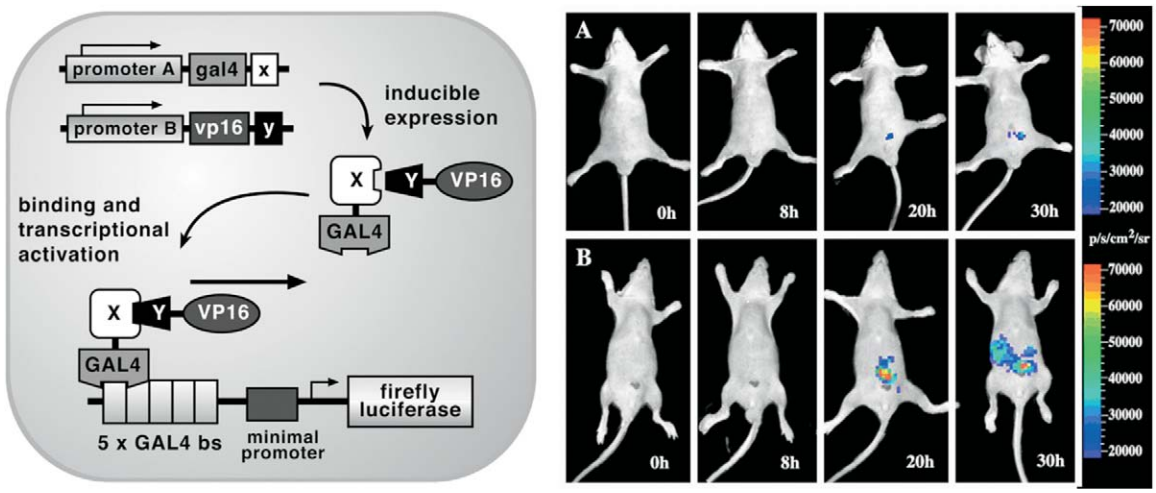


**Figure 7.** Imaging protein-protein interactions using optical bioluminescence in living mice. Proteins X and Y are chosen to be two interacting proteins. If promoter A AND promoter B are activated then bioluminescence signal is obtained due to transactivation of an optical reporter gene. The top mouse imaged for more than 30 hours shows no significant signal (control), whereas the bottom animal shows a signal due to activation of both promoters A and B. This simple AND Gate can be the basis of logical circuitry readouts for studying molecular events. Figure courtesy of Dr. Sam Gambhir, Stanford University and University of California, Los Angeles (see references 19 and 20).

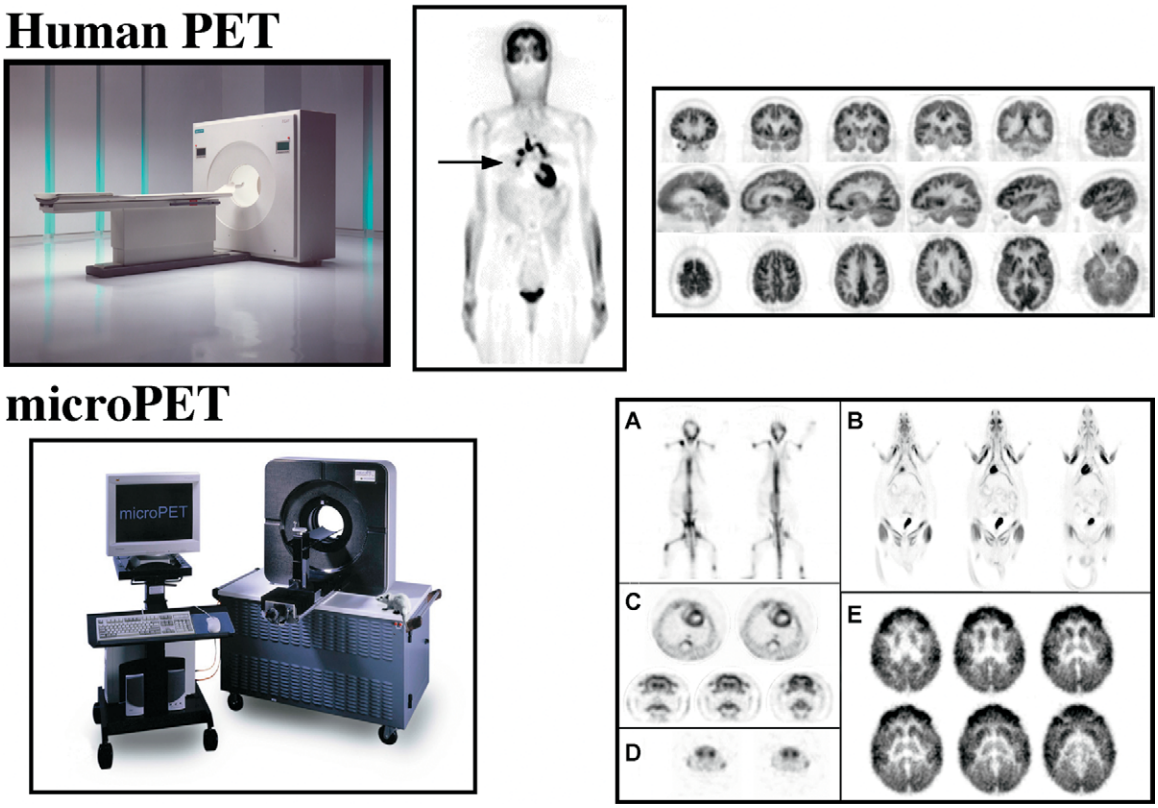


**Figure 8.** Example of imaging the same biological processes throughout the body in a patient with clinical PET scanner (CTI/Siemens, Knoxville, TN) and in a living mouse with microPET scanner (Concorde MicroSystems, Knoxville, TN). Top: Images of a single five mm thick longitudinal section through the body of glucose metabolism using glucose analog, 2-deoxy-2-[$^{18}$F]fluoro-D-glucose (FDG). The patient has metastases in the left and right lung (arrow) from a previously resected primary ovarian tumor. To the right, are tomographic images of glucose metabolism in the brain. The first, second and third rows are coronal, sagittal and horizontal planes. Bottom: microPET images of: (A) bone scan of a mouse showing two 1.5 mm thick longitudinal sections using $^{18}$F ion; (B) glucose metabolism in three 1.5-mm sections in the rat; (C) cross-sectional images through the chest showing glucose metabolism in the heart (one mm thick wall) above and in the brain below; (D) two sections in the brain of a mouse showing dopamine synthesis in the left and right striatum (12 mg each in weight); and (E) cross-sectional images of glucose metabolism in the brain of a new born monkey (brain is about 19 mm across).

Such studies can be repeated at various pharmacologic doses to assess how the pharmacokinetics throughout the body are altered by mass levels of the drug with a reference to the percent occupancy on the therapeutic target. Such studies can be performed in mice and then in patients.[27] Today, an increasing number of pharmaceutical companies have incorporated PET into their drug discovery process to be able to watch and measure their drugs and their effectiveness *in vivo* from mouse to patient. Optical imaging and MRI are also being employed to examine various structure—function relationships to aid their drug discovery process.

## Molecular Diagnostics with PET

PET has become the lead imaging technology to illustrate the general principles and value of molecular imaging of the biology of disease in patients. This is illustrated by the fact that in 20 different cancers in the diagnosis, staging, detecting recurrent disease, and assessing therapeutic responses, that on average, PET is from nine percent to 43% more accurate than conventional imaging, depending on the specific clinical question. In addition, the use of PET changes the way these patients are treated in 15% to 50% of the cases, across all these cancers and all stages of disease.[27,33]

As a second example of the value of molecular imaging in the brain, PET provides a 93% accuracy in the diagnosis of Alzheimer's disease 3.4 years before the conventional diagnosis can reach with lower accuracy the diagnosis of probable Alzheimer's.[34] MRI and CT scans of these patients are normal or show non-specific changes. Further, PET has been shown to detect Alzheimer's disease about five years before symptoms[35,36] and Huntington's disease about seven years before symptoms,[37] while in both cases the MRI and CT studies and

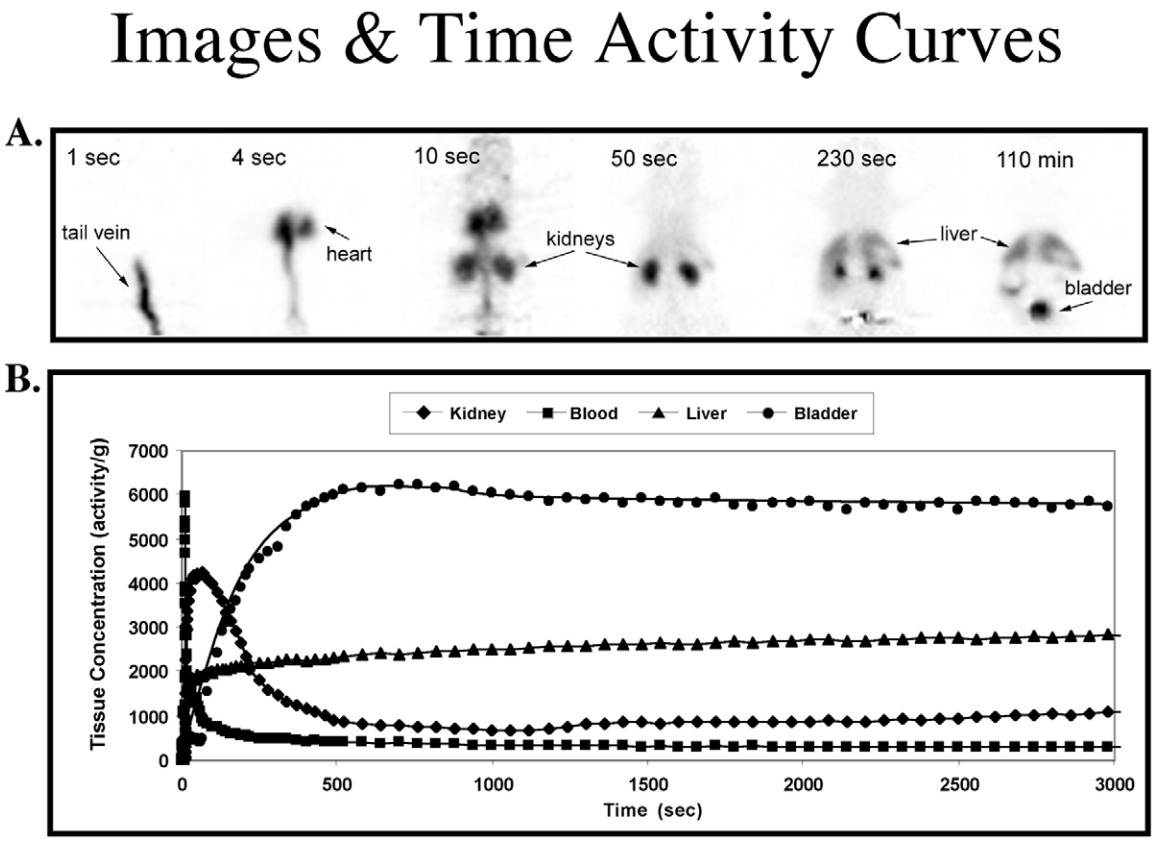


**Figure 9.** A study of pharmacokinetics of a drug analog labeled with the positron emitter $^{18}$F, fluoropenciclovir (FPCV), with a therapeutic target of herpes simplex virus thymidine kinase enzyme (HSV1-TK) in the liver. An adenovirus containing the HSV1-tk gene with a cytomegalovirus promoter was intravenously injected into the mouse to transfer the HSV1-tk gene into the cells of the liver. Greater than 95% of the virus carrier will localize in the liver. Two days after administration of the virus, FPCV was intravenously injected and a microPET scanner used to continuously record the temporal change in concentration of FPCV and its labeled metabolic products throughout the organ systems of the body. Image examples are shown at select times in A and the time activity curves for the kidneys, blood, liver, and bladder are shown in B. There is a delay in FPCV going to the bladder due to the collection and clearance time of the kidneys. Activity is retained in the liver (target) due to the phosphorylation of FPCV by HSV1-tk.

all other tests are normal. This represents the value provided from molecular imaging of the biology of disease, independent of the specific molecular imaging technology.

## Conclusion

This paper represents a tutorial of ideas, technologies and results that we believe will provide the means to revolutionize our fundamental understanding of the molecular basis of the systems biology that begins with one cell, the fertilized egg, and activates a programmed set of instructions to self-assemble an organism and maintain its normal homeostasis, as well as initiating the genetic reprogramming of cells to gain or lose functions that represent disease. From this will come new generations of informative molecular diagnostics and therapeutics. These advances will be used to form the foundation of the predictive and preventive medicine of the future.

The magnitude of the diversity of systems biology and the technologies needed to practice it effectively is staggering at this time. This is, of course, the way it is at the beginning of any paradigm shift in science, including the sequencing of the human genome and the hundreds of other genomes that are completed today. It is amazing to remember that it took about 30 years to sequence the genome of the cold virus. Because of the development of automated DNA sequencing technology it took about a day to sequence the SARS virus that has the same size genome as the cold virus. Scientific innovation and technology development can change what we believe to be possible at any given moment in time. This, however, understates the enormous effort required by scientists to achieve the paradigm shift and the benefit that results from it.

We would like to acknowledge many helpful discussions with other members of the NanoSystems Biology Alliance, including Drs. Alan Aderem, Michael Roukes, Steve Quake, Owen Witte, Charles Sawyers, and Sam Gambhir. JRH acknowledges the W.M. Keck Foundation, the Defense Advanced Research Projects Agency, the Semiconductor Research Corporation, and the Office of Naval Research for Funding. LH acknowledges funding from NIH and NSF. MEP acknowledges funding from the Biomedical & Environmental Research Division of DOE, National Cancer Institute, NIH and the Norton Simon Research Foundation.

## References

1. Lander, E.S.; Al, E. Initial sequencing and analysis of the human genome. Nature 409:860–921; 2001.
2. Venter, J.C.; Al, E. The sequence of the human genome. Science 291:1304–1351; 2001.
3. Ideker, T.; Galitski, T.; Hood, L. A new approach to decoding life: systems biology. Annu. Rev. Genomics Hum. Genet. 2:343–372; 2001.
4. Ideker, T.; Thorsson, V.; Ranish, J.A.; et al. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science 292:929–934; 2001.
5. Ideker, T.; Shannon, P.; Schwikowski, B. A particularly powerful bioinformatics package, called Cytoscape, has been developed at the Institute for Systems Biology, and has been made publicly available. Available at: http://www.cytoscape.org/.
6. Hedenfalk, I.; Ringner, M.; Ben-Dor, A.; et al. Molecular classification of familial non BRCA1/BRCA2 breast cancer. Proc. Natl. Acad. Sci. (USA) 100:2532–2537; 2003.
7. Sorlie, T.; Tibshirani, R.; Parker, J.; et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc. Natl. Acad. Sci. (USA) 100: 8418–8423; 2003.
8. Gruvberger, S.; Ringner, M.; Chen, Y.; et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. Cancer Res. 61:5979–5984; 2001.
9. Jessani, N.; Liu, Y.; Humphrey, M.; Cravatt, B.F. Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. Proc. Natl. Acad. Sci. (USA) 99:10335–10340; 2002.
10. Levsky, J.M.; Shenoy, S.M.; Pezo, R.C.; Singer, R.H. Single cell gene expression profiling. Science 297:836–840; 2002.
11. Kong, J.; Franklin, N.R.; Zhou, C.; et al. Nanotube molecular wires as chemical sensors. Science 287:622–625; 2000.
12. Star, A.; Stoddart, J.F.; Steuerman, D.; et al. Preparation and properties of polymer wrapped single-walled carbon nanotubes. Angew. Chem. Int. Engl. 40:1721–1725; 2001.
13. Cui, Y.; Qingqiao, W.; Hongkun, P.; Lieber, C.M. Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. Science 293:1289–1292; 2001.
14. Roukes, M.L.; Arlett, J.A.; Cross, M.C. Chemomechanical nanowire sensors for molecular recognition. Unpublished/Proprietary, Caltech Patent Disclosure Ref. CIT-4351.
15. Demers, L.M.; Ginger, D.S.; Park, S-J.; et al. Direct patterning of modified oligonucleotides on metals and insulators by dip-pen nanolithography. Science 296:1836–1838; 2002.
16. Lee, K-B.; Park, S-J.; Mirkin, C.A.; Smith, J.C.; Mrksich, M. Protein nanoarrays generated by dip-pen nanolithography. Science 295:1702–1705; 2002.
17. Hodneland, A.C.; Mrksich, M. Design of self-assembled monolayers that release attached groups using applied electrical potentials. Langmuir 13:6001–6003; 1997.
18. Melosh, N.A.; Boukai, A.; Diana, F.; et al. Ultrahigh-density nanowire lattices and circuits. Science 300:112–115; 2003.
19. Heath, J.R.; Ratner, M.A. Molecular electronics. Physics Today 56:43–49; 2003.
20. Unger, M.A.; Chou, H.P.; Thorsen, T.; Scherer, A.; Quake, S.R. Monolithic microfabricated valves and pumps by multilayer soft lithography. Science 288:113–116; 2000.
21. Thorsen, T.; Maerkl, S.J.; Quake, S.R. Microfluidic large scale integration. Science 298:580–584; 2002.
22. Fu, A.Y.; Chou, H.P.; Spence, C.; Arnold, F.H.; Quake, S.R. An integrated microfabricated cell sorter. Anal. Chem. 74:2451–2457; 2002.
23. Braslavsky, I.; Hebert, B.; Kartalov, E.; Quake, S.R. Sequence information can be obtained from single DNA molecules. Proc. Natl. Acad. Sci. (USA) 100:3960–3964; 2003.
24. Lippincott-Schwartz, J.; Patterson, G.H. Development and use of fluorescent protein markers in living cells. Science 300:87–91; 2003.
25. Contag, P.R.; Olomu, I.N.; Stevenson, D.K.; Contag, C.H. Bioluminescent indicators in living mammals. Nat. Med. 4:245–247; 1998.
26. Massoud, T.; Gambhir, S.S. Molecular imaging in living subjects: seeing fundamental biological processes in a new light. Genes Dev. 5:545–580; 2003.
27. Cherry, S.R.; Shao, Y.; Silverman, R.W.; et al. MicroPET: a high resolution PET scanner for imaging small animals. IEEE Trans. Nucl. Sci. 44:1109–1113; 1997.
28. Chatziioannou, A.F.; Cherry, S.R.; Shao, Y.; et al. Performance evaluation of microPET: a high resolution lutetium oxyorthosilicate PET scanner for animal imaging. J. Nucl. Med. 40:1164–1175; 1999.
29. Phelps, M.E. Positron emission tomography provides molecular imaging of biological processes. Proc. Natl. Acad. Sci. (USA) 97:9226–9233; 2000.
30. Ray, P.; Pimenta, H.; Paulmurugan, R.; et al. Noninvasive quantitative imaging of protein-protein interactions in living subjects. Proc. Natl. Acad. Sci. (USA) 99:3105–3110; 2002.
31. Paulmurugan, P.; Umezawa, Y.; Gambhir, S.S. Noninvasive imaging of protein-protein interactions in living subjects by reporter protein complementation and reconstitution strategies. Proc. Natl. Acad. Sci. (USA) 99:5608–5613; 2002.
32. Tjuvajev, J.G.; Finn, R.; Watanabe, K.; et al. Non-invasive imaging of herpes virus thymidine kinase gene transfer and expression: a potential method for monitoring clinical gene therapy. Cancer Res. 58:4333–4341; 1998.
33. Shields, S.F.; Grierson, J.R.; Dohmen, B.M.; et al. Imaging proliferation in vivo with F 18FLT and positron emission tomography. Nat. Med. 4:1334–1336; 1998.

34. Gambhir, S.S.; Barrio, J.R.; Phelps, M.E.; et al. Imaging adenoviral-directed reporter gene expression in living animals with positron emission tomography. Proc. Natl. Acad. Sci. (USA) 96:2333–2338; 1999.
35. Gambhir, S.S.; Czernin, J.; Schwimmer, J.; Silverman, D.H.; Coleman, R.E.; Phelps, M.E. A tabulated summary of the FDG PET literature. *J Nucl Med* 42 (Supplement): 1S–93; 2001.
36. Silverman, D.H.; Small, G.W.; Chang, C.Y.; et al. Positron emission tomography in evaluation of dementia: regional brain metabolism and long-term outcome. JAMA 17: 2120–2127; 2001.
37. Small, G.W.; Mazziotta, J.C.; Collins, M.T.; et al. Apolipoprotein E type 4 allele and cerebral glucose metabolism in relatives at risk for familial Alzheimer disease. JAMA 12:942–947; 1995.