Physical Sciences-Based Frontiers in Oncology

THE CODING, DECODING, TRANSFER, AND TRANSLATION OF INFORMATION IN CANCER

October 29-31, 2008

Meeting Report

NATIONAL CANCER INSTITUTE

NATIONAL INSTITUTES OF HEALTH

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA

The Ritz-Carlton • Pentagon City • Arlington, VA

Cover image courtesy of NISE Network, Viz Lab, and Linda Nye.

Physical Sciences-Based Frontiers in Oncology

THE CODING, DECODING, TRANSFER, AND TRANSLATION OF INFORMATION IN CANCER

October 29-31, 2008

Meeting Report

The Ritz-Carlton • Pentagon City • Arlington, VA

Contents

Executive Summary

"The Coding, Decoding, Transfer, and Translation of Information in Cancer" is the third in a series of 2008 think tanks convened by the National Cancer Institute (NCI) to explore innovative ideas, concepts, and theories from the physical sciences that could inform and enable a fundamental understanding of cancer at all scales. The prior two meetings, "Integrating and Leveraging the Physical Sciences To Open a New Frontier in Oncology," held February 26-28, 2008, and "A New Look at Evolution and Evolutionary Theory in Cancer," held July 13-15, 2008, engaged over 200 experts from physics, mathematics, physical chemistry, and basic and clinical cancer research. Outcomes from the first think tank identified four convergent themes of critical importance to cancer research: the "physics" of cancer (the forces, thermodynamics, gradients, etc. that govern behavior at all scales); the role of evolution and evolutionary theory in cancer; information flow, translation, and information theory in cancer; and "deconvoluting" the complexity of the disease. The second think tank explored the potential value of studying cancer from an evolutionary perspective and further highlighted the pressing need to consider questions related to information sources, flow, and contextual translation in cancer at all scales (molecules, organelles/cells, tissues, organisms). It was the consensus of the first two meetings that the complex processes that drive the emergence of the malignant phenotype in cancer were information rich, and, like evolution, these areas of science offered significant opportunities to better understand and control cancer.

The current meeting, "The Coding, Decoding, Transfer, and Translation of Information in Cancer," was designed to discuss the wide range of topics that constitute these fields. The meeting, a facilitated think tank, included a few broad keynote presentations to introduce major topic areas; panel discussions that pursued specific research areas and findings; and brainstorming sessions that included all of the participants. In addition, smaller working groups considered a number of the major questions or "grand challenges" surrounding the coding, decoding, transfer, and translation of information in cancer from the standpoint of transdisciplinary research and associated resource needs.

Specifically, the meeting comprised a conceptual "arc" that began with a "stage setting" presentation by **Dr. Robert Phillips**, California Institute of Technology (Caltech), which focused on information management and the nature of cellular decision-making. In offering his perspectives, **Dr. Phillips** emphasized the usefulness of physical measurements and the quantitative analysis of biological systems as bases to provide context for understanding the role of information and its translation and measurement in cancer biology. **Dr. John Niederhuber**, Director of the National Cancer Institute, provided context for the meeting by giving a brief overview of the current state of cancer research and identifying some of the key knowledge gaps that drove the design of the current meeting. **Dr. Christoph Adami**, Caltech, gave a keynote presentation on information theory and its potential value in understanding information flow in cancer, particularly the use of sequence information and implications at the level of mutated genes. The meeting proceeded to consider the nature of the "information" in cancer with keynote presentations by **Dr. David Haussler**, University of California, Santa Cruz, and **Dr. Phillip Sharp**, Massachusetts Institute of Technology. Dr. Haussler discussed our current understanding of DNA, genes, transcription, and information translation across time. **Dr. Sharp** explored the small RNAs and their role in regulating information flow in cancer. Both of these speakers emphasized the value of comparative genomics and the promise of the transcriptome to uncover functional elements in cancer.

In a subsequent discussion, panelists considered a range of topics related to cell signaling, cellular decision-making, and the translation of information in cancer—with significant consideration of the spatial and microenvironmental contexts. A second panel considered questions of contextual translation of information in cancer from the standpoint of how malignant phenotypes evolve, with specific emphasis on the physics of these processes. All of the meeting panels examined the multiscale, temporal, and spatial nature of information transfer (from germline to tissue and organism levels),

differences and similarities between normal and cancer information, and tools being used to decipher information and processes.

The meeting moved to consider the next rational question in the conceptual "arc": Is an understanding of the nature of the information per se and mechanisms of its transfer and contextual translation a rational basis for altering the progress of cancer? This question was explored in depth by **Dr. Daniel Hillis**, Applied Minds, Inc. He discussed cancer as an "emergent complex system" and considered strategies for control at the patient level by exploring what constitutes and drives emergent systems from an information standpoint.

During the course of the meeting, barriers (grand challenges) were identified that limit development of the complex field of information management at all scales in cancer. The participants worked in four small groups to reach consensus on new directions and focus areas for research, needed tools and technologies, and other resources needed to address research requirements and build a new transdisciplinary field of information coding, decoding, transfer, and translation in cancer. The four groups were (1) major overarching research questions, (2) nature of the critical information in cancer, (3) communication in cancer at multiple scales, and (4) technology, models, and tools. The outcomes from the four groups are presented in detail in the report that follows.

In summary, this meeting considered the critical topic of information in cancer in the context of both the biological and physical sciences—what it is, how it is transferred, and how it is translated. Examples of several important concepts that emerged from this meeting include the following: the "gene" can no longer be viewed as a single entity but instead as a complex information coding construct; going beyond a genocentric view of cancer to measure state changes in cancer will be very important; information in cancer is context dependent and scale specific (e.g., cells, tissues, whole organisms); information and its management in cancer must be considered across space and time; and cellular architecture and measuring communications through structural pathways are important in understanding contextual translation. Although cancer will be defined by large amounts of information at different scales, these detailed datasets may not reflect the level at which cancer is best controlled. From an information standpoint, cancer is an emergent complex system, and models of these types of systems suggest that their control is often not at the level where the amounts and types of information seem most compelling.

Session Summaries

The Coding, Decoding, Transfer, and Translation of Information in Cancer

Day 1: Wednesday, October 29, 2008

Meeting Background and Introductions

Anna D. Barker, Ph.D., Deputy Director, NCI

Dr. Barker greeted the attendees and presented an overview of the general objectives of this series of think tanks. These meetings are held to explore various areas of the physical sciences that are critical in developing both a fundamental understanding of cancer and new strategies for cancer control.

This is the third meeting in a series focused on applying new thinking from the physical sciences to examine major questions in cancer, often in seemingly unorthodox ways. The first exploratory meeting, *Integrating and Leveraging the Physical Sciences To Open a New Frontier in Oncology,* was held February 26-28, 2008. Four overarching themes emerged at that meeting for further exploration: (1) the "physics" of cancer (forces and mechanics, thermodynamics, gradients, etc.); (2) evolution and evolutionary theory in cancer; (3) information coding, transfer, translation, and information theory in cancer; and (4) the complexity of cancer. The second think tank, *A New Look at Evolution and Evolutionary Theory* in Cancer, July 13-15, 2008, identified major research questions and challenges that, if addressed, could significantly improve our understanding of cancer. The major input from this meeting was that value could be gained by placing what we know about cancer into an evolutionary framework and using this framework to provide future direction for cancer research. The third meeting stemmed from many of the conversations at the first two meetings, where questions were raised on the role of all aspects of coding, decoding, and translation of information and information theory in understanding evolution of cancer as an integrative, complex, and emergent complex system.

Dr. Barker introduced Dr. Niederhuber, Director of the NCI, who introduced the first keynote presenter.

Welcome and Introduction of Keynote Presenter

John E. Niederhuber, M.D., Director, NCI

Dr. Niederhuber greeted the attendees, echoing Dr. Barker's comments about the value of this series of meetings to date, and introduced Dr. Robert Phillips. Dr. Phillips is Professor of Applied Physics and Mechanical Engineering at the California Institute of Technology in Pasadena, California. Dr. Phillips' group works on physical biology of the cell, physics of genome management, and use of physical models to explore biological phenomena. He is coauthor of the soon-to-be-published textbook *Physical Biology of the Cell*. Dr. Niederhuber noted that Dr. Phillips is a self-described lifelong student of the scientific approach to understanding nature and the engineering basis for controlling it, which, he pointed out, represented an excellent rationale for all of these think tanks.

Keynote Presentation

Is DNA a Molecule? Musings on Good Cells Making Bad Choices

Robert Phillips, Ph.D., Professor of Applied Physics and Mechanical Engineering, Division of Engineering and Applied Science, California Institute of Technology

Presentation Highlights (For a full graphical representation of this talk, see Figure 1, Appendix 1.)

- Historical lessons from successful solution of biological problems are that physical analysis of biological questions requires:
  - Detailed quantitative analysis, employing mathematical approaches and tedious measurements.
  - A requirement for theoretical structures to elucidate the data.
  - In addition to the advances in biology and physics, payoffs in advances in enabling technology occur in applied sciences, health, and energy disciplines, and support for the development of technological tools is critical.
- Quantitative analysis of cellular decision-making contributes to the understanding of genome management.
  - Control of nucleic acids, proteins, and the processes connecting them requires constant information management.
- Predictive frameworks are needed for the increasing amount of biological data being generated on functional relationships.
- The new generation of life scientists should be educated in the use of quantitative analysis with data.

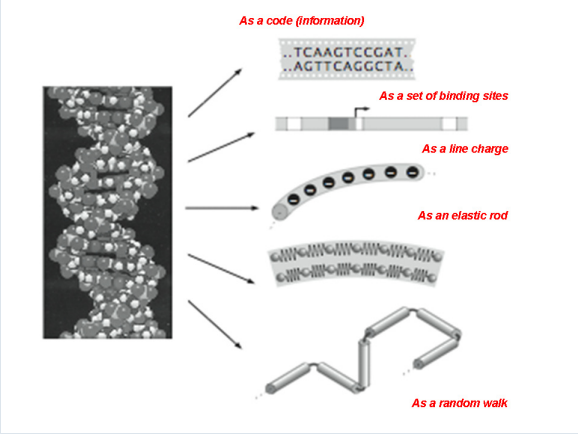
Beginning with his general philosophy, Dr. Phillips suggested that a productive path to progress in biology and allied engineering disciplines lies in the detailed physical dissection of biological problems employing the interchange of theory and experimentation, including development of new enabling technologies. He then discussed "the big picture" based on a series of hypotheses, illustrated with historical examples that summarized progress at the interface between physics and biology, with possible consequences for health, energy, and cancer. Dr. Phillips emphasized that the quantitative dissection of biological problems will continue to yield new insights into both biology and physics, with clear benefits in areas such as human health. Historically, progress has come from detailed case studies; mathematical approaches and tedious measurements have helped solve biological questions and have led to great discoveries. For example, the work of Mendel, Morgan, and Sturtevant used frequency counting to begin the process of defining genes and mapping chromosomes. Broad principles can be generated from studying specific problems and understanding specific examples. Thus, for example, detailed quantitative dissection of cellular decision-making can contribute to understanding genome management and mismanagement. However, data by themselves are insufficient. It is critically important to have theoretical structures to partner with measurements. As contemporary biological studies become increasingly data rich, the demand is clear and is increasing for biological theory to provide a needed framework to understand the data.

Significant technological advances in the past have come unexpectedly from those pursuing strictly scientific agendas (e.g., the Curies' discovery of radioactivity and the huge impact of radioactivity on subsequent experimentation and discovery). Support for the development and use of technologies and approaches such as synthetic biology to enable measurements of systems will continue to be important. Dr. Phillips also pointed out that follow-on use of engineering disciplines, or applied science, works best as a rational outgrowth of intellectual infrastructure, not as enlightened empiricism. The development by Roger Tsien's group of fluorescent proteins as tools for cell biology is one example of this.

Dr. Phillips discussed the study of cellular decision-making from the standpoint of physical model building and quantitative experimentation as being illustrative of normal and cancer cell information processes. Francis Crick referred to nucleic acids and proteins as the "two great polymer languages," and the processes connecting them require constant cellular decision-making or genome management for meaningful exploitation of the sequences. Examples of "good" cellular decision-making include the *lac* operon (as described by Jacob and Monod[17]) and the development of an embryo into a multicellular individual.[5]

Cellular decision-making also links the informational and physical characteristics of genomic DNA; understanding the information (and its corruption) requires understanding its physical manipulations. In considering DNA as a molecule, Dr. Phillips pointed out the focus on DNA as a series of letters, as information and argued using illustration with case studies that physical implementation of the molecule has everything to do with readout of DNA content information. For example, DNA can be

understood as a code (information), a set of binding sites, a line charge, an elastic rod, or a random walk.



Physical manipulations are required for information acquisition and understanding and often for information corruption. Experimentation that elucidates the interplay between the informational and physical characteristics of genomes has illustrated the importance of chromosomal DNA organization, nucleosome positioning, DNA packaging with regard to gene expression and state (including methylation, etc.), affinity of binding to sites, and dependence on packaging depth.[1-4]

Cellular decision-making also involves signaling pathways or networks. Experimentally and quantitatively, there is a need to determine the number of components, locations, and timeframes. For instance, to dissect a network quantitatively, one can systematically vary parameters and examine the biological outcome. Estimates can also be useful. Dr. Phillips added that it might be necessary and important to find new technical methods to conduct measurements.

Dr. Phillips concluded by noting that a new generation of life scientists is needed that uses quantitative analysis and data as part of the normal toolkit: "Biological data have forced this issue—if people are going to go to all the trouble of making and presenting quantitative measurements, the intellectual response to those data needs itself to be quantitative."[6]

Discussion Highlights: Two major areas of discussion followed Dr. Phillip's presentation. In a discussion of how physics can be used to understand the biological processes built through evolution, Dr. Phillips noted that biological systems have to respect the laws of physics. He also pointed out that there is not a proper appreciation for the use of theory in biology and that even wrong models can be useful.

Think Tank Process and Outcomes Overview: Dr. Barker introduced the think tank Facilitator, Mr. Robert Mittman, who has served this role for all of the meetings. Mr. Mittman gave an overview of the process for the think tank and briefly discussed expected outcomes. He further explained that the process would be described in detail on Day 2 of the think tank, when all participants would be on hand.

Day 2: Thursday, October 30, 2008

NCI's Physical Sciences-Based Frontiers in Oncology Series

Think Tank Process and Outcomes Overview

Anna D. Barker, Ph.D., NCI, and Robert Mittman, M.S., M.P.P., Chairman, Facilitation, Foresight, Strategy

Dr. Barker briefly reviewed prior think tanks for the new arrivals and introduced the facilitator, **Mr. Robert Mittman**. Mr. Mittman added to Dr. Barker's introduction by describing the current meeting as organized into four conceptual segments to reflect the four central questions posed. The conceptual segments are also designed to set the stage for achieving NCI's desired outcomes for the meeting, that is, development of innovative strategies, models, and approaches to help build a transdisciplinary field of cancer information coding, decoding, transfer, and translation science, as well as a theoretical foundation for this complex process.

Some central goals that derive from the questions presented in the agenda were posed for consideration in terms of the conceptual meeting framework, including:

- Identification of the range of information sources and processes in cancer biology at different length and time scales
- Exploration of major research questions, future strategies, and coherent theoretical approaches that will enable a fundamental understanding of cancer across these scales
- Identification of the barriers that limit timely progress in the field
- Given that progress is achieved in the first three, provision of input and guidance in structuring and prioritizing research questions for NCI and individual investigators (e.g., research strategies, data management approaches, infrastructure, etc., to support and inform accomplishment of research goals)

To begin the scientific program, Dr. Barker introduced **Dr. John E. Niederhuber**. As many of the scientists on hand were not from the field of cancer research, Dr. Niederhuber summarized the state of the science in cancer research as he did at the initial meeting in this series; his remarks focused on current trends and concepts in cancer research from his perspective as a surgeon with interests in stem cell research and crosstalk in the microenvironment.

Welcome and Keynote Presentation

State of the Science in Cancer Research

John E. Niederhuber, M.D., Director, National Cancer Institute, National Institutes of Health

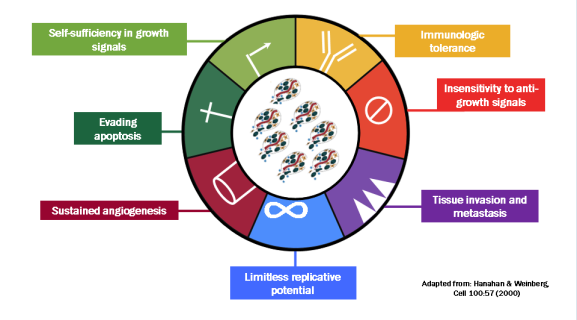
Presentation Highlights (For a full graphical representation of this talk, see Figure 2, Appendix 1.)

- Entering an unprecedented era of discovery and a new era of medicine.
- The challenge will be in understanding the **complexity of the regulatory systems** and the **information** that drives it
  - Tumor microenvironment, "niche"—need to control the microenvironment.
    Dynamics of cellular communication, chemical gradients.
    Effect on cancer cells and receptiveness to the process of cancer spread.
  - Tumor cell heterogeneity and "cancer stem cells"—need to understand the role of cancer stem cells.
    The power of self-renewal, travel to other tissues, capacity as progenitor cells.
  - Need to understand the structural organization of information—spatial position may be a diagnostic tool.
  - New levels of imaging reveal new dimensions of complexity.
- Requirement for transdisciplinary research teams.

Dr. Niederhuber's opening statement reflected the overriding reason to hold these think tanks. The significant human and economic burdens of cancer create a critical need to address major barriers. Progress has been made, as evidenced by declining death rates in certain tumors, due primarily to early diagnosis, fewer smokers, and use of vaccines. He reemphasized that the focus for this meeting is to explore what physics, physical chemistry, and applied mathematics can bring to cancer biology and determine how this group of scientists can most effectively become involved in further advancing cancer research.

Given that cancer is a disease of genes and altered genes, the power to sequence the human genome has ushered in an unprecedented era of discovery and transformation of medicine. However, it is not enough to just sequence the code. **The real challenge will be to understand the complexity of the regulatory system and the information that drives it** (not just epigenetic modifications such as methylation but also complexity at the structural level). The huge amount of information that is the complex genetic code is translated into functional changes in the cell, and these complex changes result in cell transformation, with the numerous phenotypic changes we see in cancer. While the focus of cancer research has been on tumor cells for many years, we increasingly recognize that the tumor is not just disease of abnormal genes but also a process by which cancer growing in its own microenvironment can invade and metastasize. A tumor is a complex "organ," and the focus of cancer research is shifting from tumor

cells to understanding the increasing complexity of the organ system of cancer growth.



The complexity of the tumor microenvironment and the tumor cell environmental "niche." Cells are not autonomous; the microenvironment is important. The complexity of a tumor includes dynamic communication processes that drive chemical gradients and other effector mechanisms that occur among heterogeneous cell types that populate the tumor microenvironment. Factors produced by these aberrant cells—for example, growth factors, chemokines and chemotactic factors, and proteases—can alter aspects of tumor cell behavior and are part of the process of metastasis and invasion. What comes first, changes/abnormalities in the microenvironment or in the cells? The environment may have to be set for genetic changes in cells to be recognized and implemented.

Increasingly, it is clear that cell migration does not occur by chance but that it is a complex process. Knowing the cell's environmental niche is important. Investigations into creation of a receptive, premetastatic environment include consideration of fibronectin deposition, migration of endothelial progenitor cells, vascular organization, and other factors.[7] In experimental models, a tumor cell migrating into a normal microenvironment grows normally, but if the cell migrates into a supportive abnormal niche, it displays cancerous properties. Dr. Niederhuber predicted that **to control cancer as a disease, it will be necessary not only to operate at the cancer cell level; there also will be a real need to control the microenvironment.**

Tumor cell heterogeneity and "cancer stem cells." The complexity of a tumor is also characterized by the heterogeneity of the tumor cell population. Within tumors, a small number of cells demonstrate unusual characteristics, including self-renewal capacity or stem-cell-like properties, and are referred to by some as "cancer stem cells." **Understanding the role of cancer stem cells will be critical to developing a complete picture of cancer.** It is unclear what mechanism allows cancer to return as metastatic disease after a patient has been free from cancer for many years. What if therapies (both chemotherapy and radiation) effectively treat most tumor cells but are not effective at treating the small percentage of differently programmed cells, ultimately resulting in recurrence? If tumor treatments target stem cells, can more differentiated cells, which are programmed to die, then be more easily eliminated? Another key question is whether the process of cancer initiation takes place in the stem cell or in a progenitor cell. For example, genetic changes could occur in a progenitor cell that reprograms a cell to be more like a stem cell.

The complexity of the structural organization of the information. Mitochondrial and nuclear structural organization are critical areas for future research. NCI imaging studies of genome organization in three-dimensional (3D) space demonstrate that chromosomes are not randomly positioned. Definitive patterns have been measured in normal vs. breast cancer cells. This work illustrates that regulatory processes are involved in structural organization of the chromosomes and that the gene position is not random. **In fact, spatial position may be useful as a diagnostic tool to differentiate normal from premetastatic and malignant cells and tumor types.**

In conclusion, cancer is a complex disease, and there may not be a more complex problem than metastatic cancer. The levels of complexity include protein-protein interactions, chemical gradients, energy-time interactions at the target, and as-yet unexplored physical forces that are important to understanding migration of cells, cell changes, and forces involved in changing the environment. There has never been a more exciting time to work in science. The rapidly developing technologies that drive complex research require that science of the future involve teams coming together to solve problems. Dr. Niederhuber added that what we learn in studying cancer will inform the diagnosis and treatment of other diseases.

Information Theory in Molecular Biology: Key to Understanding Information Transfer, Signaling, and Translation in Cancer

Christoph C. Adami, Ph.D., Professor, California Institute of Technology

Presentation Highlights (For a full graphical representation of this talk, see Figure 3, Appendix 1.)

- Predictions can be made about a system with accuracy better than chance.
  - Quantifies the amount of information in messages.
  - Quantifies the capacity of channels to transmit information (given noise).
- Information is essentially contextual. Changes in the environment (niche) result in changes in the information.
- Fitness depends on information about the environment.
  - Cells and organisms use information in genes for survival.
  - Fitness changes imply changes in information content.
- Shannon's entropy: mechanism to quantify the probability of correctly predicting the state of X.
  - The information stored in a gene is the difference between the maximal and actual entropy.

Dr. Adami presented an overview of information theory, emphasizing characteristics of the theory that are potentially useful in cancer research. Information theory was developed at least 50 years ago, pioneered by Claude Shannon; because it can be used to simplify complicated problems, it has been established as a generally applicable tool for understanding complex systems.

Information theory can be viewed as a form of nonequilibrium statistical physics. More generally, **information theory examines the relative state of the detectors**.

- Information theory allows the information keeper to make **predictions about a system** with better-than-chance accuracy. For example, information theory can predict residue at a specific site using sequence information.
- The theory makes **information by its very essence a contextual quantity**, a key concept.
- **The system is important**. The information is dependent on the system; if the system changes, it is no longer information. In the example above, the sequence stored in a genome is in the context of the environment in which the organism lives. The organism similarly makes predictions about its environment; this environment (i.e., the niche) is very important in determining what the information essentially is.
- **There is a connection between fitness and information**, another key concept. As in evolution, fitness permits an organism to live. The more information available about the environment, the better the chance for survival in the environment. Fitness is a long-term predictor about the success of a gene.
- The theory quantifies the amount of information in messages and **provides the means to quantify the capacity of channels to transmit the information**. Note that information can be distributed among many agents.

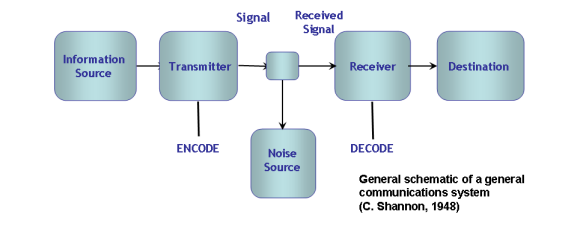
Shannon's formula defines the entropy, H, of a random variable or molecule, X, as a sum over the set of probabilities, $p_1 \ldots . p_N$, of the possible states of X, x_i.

$$H(X) = - \sum_{i=1}^{N} p_i \log p_i$$

Shannon's entropy provides a means to quantify the probability of correctly predicting the state of X. If the entropy or uncertainty is very large, the probabilities will be very small. (The uncertainty is how much is not known about something.). If the entropy is 0 (i.e., everything is known about it), the probability will be 1.

For x_i molecules in pools that are functionally the same, the actual entropy of the pool is much less than the maximal entropy, and the **difference between the actual entropy and the maximal entropy is the information in the genes**. The actual measured entropy is conditional, as it is dependent on the environment. For example, one can measure the information stored in genes by examining the difference between maximal entropy stored in genes by a set of molecules and the actual measured conditional entropy within the environment of the molecule. Thus, for a 100-amino acid protein, the maximal entropy per site is 1, and the maximal entropy is 100. The actual measured entropy will be smaller. (The entropy of our DNA is very, very small. Our DNA is very similar, with the exception of the single nucleotide polymorphisms.)

Dr. Adami proposed that fitness and information are linearly related; if w is fitness, $I \approx k \log w$. An example is the study of the evolution of drug resistance in HIV. Due to a rapid mutation rate, the HIV virus can adapt quickly to a new environment. In general, a loss of information is observed with the development of drug resistance, due to the accumulating mutations, some of which are compensatory. To test this, the (loss or gain of) information content of the HIV-1 protease (a 99 mer) can be calculated by using the mutation/substitution probabilities at each residue. If mutations between residues are not correlated, the entropy of the protease can be found from the sum of the entropies of each residue, using substitution probability at each residue. Using sequence information from the Stanford HIV drug-resistance database, substitution probabilities per site were calculated, and changes per site over time were used to obtain a profile of the entropy. (Using the database, changes over time can also be examined.) Using normalized entropies of 0–1, 0 is the entropy if only one amino acid is found per site. A value of 1 represents a case of the same probability for each of the amino acids (1/20). Thus, the total information, or entropy per site, is 1 entropy per site, and if not correlated, the entropy would be the sum of the information per site. Results showed that four areas in the 99 mer were found to have low entropy (high information). These areas of the sequence are where most of the information is coded. When analysis is corrected for correlated mutations between sites, the corrected analysis indicates that as time goes on and the environment becomes more complex, treatment with multiple drugs actually creates more information-rich viruses rather than fewer.



General schematic of a general communications system (C. Shannon, 1948)

Shannon's theory also quantifies the amount of information that can be sent across a channel with the accuracy of the channel (given noise) and a decoder.

There are two ways of looking at information transmission across channels in molecular biology. The first is transmission of information across generations in evolution. The second is the transmission of information from the environment to the cell machinery (i.e., the information processing capacity of a cell). The channel view monitors information processing at the single-cell level. As one example, an artificial cell model developed in **Dr. Adami's** laboratory is being used to study information transmission pathways. Enzymes, chromosomes, transcription factors, membrane proteins, etc. are all examples of information transmission channels. It may be possible to measure the capacity of these channels. For example, if a network view of interacting proteins is used, the relationship between a pair of proteins can be determined by measuring the output from protein 2 after modifying the input to protein 1. If there is no change in protein 2 due to a change in 1, then there is no correlation between the proteins and the capacity is 0 for both cases. Measurements are repeated for the rest of the protein pairs in the network; some pairs will demonstrate a clear correlation. Measurement of the correlation across the pairs develops the channel relationships' network picture and capacity.

In conclusion, **Dr. Adami** pointed out that most of what he discussed in relation to cancer is based on the assumption that if cancer is a disease in which single cells with a mutated genome gain a replicative advantage over other cells, then information theory is a general tool to study cancer genes—because fitness changes imply changes in information content. It may not be possible to measure the fitness change of a particular gene, but if sequence data are available, it might be possible to measure changes in the information content. Because information may be used as a proxy for fitness, it can be used to reveal the association between oncogenes and tumor suppressor genes. One can also use the theory to characterize information transmission channels that can lead to a better understanding of changes in signal transduction.

Discussion Highlights: A key discussion point was how to quantify transmission of information in a noisy channel, for example, collections of cells. **Dr. Adami** pointed out that noise in the channel can be obtained from measuring the relationship between input and output. Imagine the relationship between input and output signals, for example, the *lac* operon. Gene activity is dependent on lactose in the environment—high lactose: gene on; low lactose: gene off. A plot can be used to calculate the channel capacity (1 bit). Any signal with lactose absent is the noise (the low-level activity).

Reading Information in the Germline and Cancer Genomes by Its Evolutionary Signature

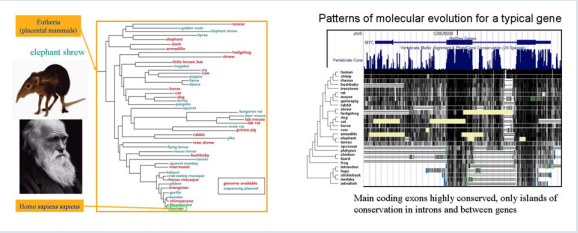
David Haussler, Ph.D., M.S., Professor, University of California, Santa Cruz

Presentation Highlights (For a full graphical representation of this talk, see Figure 4, Appendix 1.)

- The germline genome is structured by evolution; similar structural changes occur in the cancer genome (chromosomal rearrangements, duplications, deletions, mutations).
- Sequence alignments, using comparative genomics, lead to identification of selection coefficients used to identify patterns of coding and noncoding regions across evolution.
- Comparative genomics is being used to identify critical changes and important functional elements in cancer (gene expression and pathways).

The germline genome changes that occur during evolution due to chromosomal rearrangements, deletions, additions, and single point mutations are also present in the cancer genome. In his presentation, **Dr. Haussler** discussed parallels between genomic physical changes that drive evolution on a population basis and those that give rise to cancer and drive its progression. He further pointed out that the power of comparative genomics can be used to identify critical changes and important functional elements in cancer.

Deletions, amplifications, and single base changes result in structural changes that not only give rise to either the creation or loss of germline genes in evolution but also change gene expression, inactivate genes, and disrupt interacting pathways in cancer. Single base changes that result in inactivation of the p53 gene (21,588 somatic mutations cataloged; 15,387, or 71%, are missense mutations) are located in the core domain for DNA binding. For example, consistent tissue-specific patterns of amplifications and deletions were reported in breast and brain tumors in relation to normal tissue.[9] In addition, somatic and germline nonsilent mutations, amplifications, and deletions have been found in brain tumor signaling pathways (elevations in p53, RB and receptor tyrosine kinase/Ras/phosphoinositide-3-kinase pathways compared with control tissue).



Comparative genomics is currently employed to map out the evolutionary history of the genome. This information then can be used to identify regions of the genome that could be critical for adaptive events. In reconstructing the past 100 million years of evolution, key events or sequences that gave rise to mammals were identified in regions where many changes occurred as well as in regions with highly conserved coding exons, points of introduction of new introns, etc. When sequences are aligned, **selection coefficients, or entropy**, can be measured. Interesting patterns of selection coefficients are being uncovered where the patterns distinguish sequences that do or do not function as coding regions. Similarly, comparative genomics can be used to better understand cancers. While mutations and changes in gene expression have been demonstrated between normal and tumor cells, the information can be amplified by examining pathways. Combining the information about changes in copy number in somatic cells and germline cells provides the statistical power needed to determine whether a pathway is important in the development of a cancer.

Some of the lessons learned from patterns of molecular evolution for a typical gene:

- **Main coding exons are highly conserved**, while only islands of conservation occur in introns and between genes.
- **Neutral drift is defined as a genetic change that does not affect the organism**. Mutations frequently occur in protein-coding regions; some do not alter the protein and thus do not affect the fitness of the organism—for instance, a change in the third DNA base in a codon.
- **Negative selection is rejection of a change that decreases fitness**. Mutations that would change the protein, thereby reducing fitness, are rejected by natural selection, and the DNA is conserved. This results in a pattern of selection that identifies coding DNA.
- **Positive selection is a genetic change, or mutation, that increases fitness.**
- There are ~500,000 conserved noncoding regions in the human genome, some more conserved than others; these regions extend over hundreds of bases and cluster within

~1 mb of developmental genes. Sites in these regions exhibit strong selective pressure, with selection coefficients three times higher than coding regions. Furthermore, some noncoding regions have switched from negative to positive selection.

- The evolution of vertebrates was greatly facilitated by transposons derived from viruses. Most of the genome consists of molecular "fossils" of transposons, mobile DNA from defective viruses, and turnover of noncoding DNA, largely from the activity of transposons. Many conserved noncoding elements derive from ancient transposons. Comparative analysis with the opossum genome showed that at least 15% of the conserved noncoding elements specific to placental mammals came from known transposons. Interestingly, ChIP data on binding sites for human p53 indicate that one-third are primate specific and derived from two families of endogenous retroviruses.

In conclusion, research to date demonstrates that while much of the information in the genome is noncoding regulatory information, there is limited information on the coding information. However, with enough data and comparative genomics, the important functional elements can be recognized by their patterns of selection, in both germline and tumors. Finally, **Dr. Haussler** cautioned that we should expect the unexpected when looking for the origins of functional elements in the genome.

Keynote Presentation

The Rest of the Story: The Small RNAs and Cancer

Phillip A. Sharp, Ph.D., Professor, Massachusetts Institute of Technology

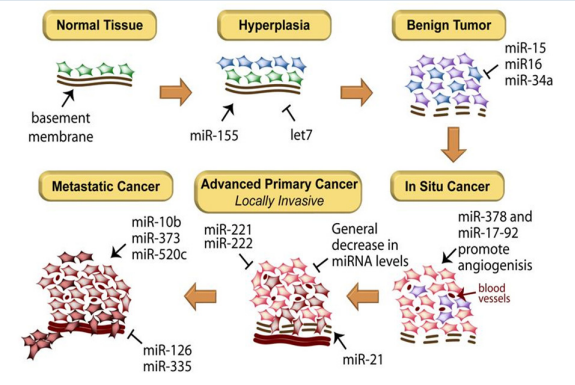
Presentation Highlights (For a full graphical representation of this talk, see Figure 5, Appendix 1.)

- Importance of microRNAs (miRNA) in regulation of biological pathways:
  - Target-specific mRNA regulates protein expression of up to 50% of all genes in vertebrates.
  - mRNA sequences enriched in complementary sequences to miRNAs; 87 evolutionarily conserved seed families of miRNAs.
- Loss of miRNA regulation has been correlated with cancer progression.
  - Changes in specific mRNA molecules have been identified in cancers.
- An estimated 94% of human genes (multiexon genes) undergo alternative splicing, some tissue- and context-specific.
- Evidence of coregulation of splicing and polyA cleavage—a mechanism to coordinate the ORFeome with the UTRome?

Dr. Sharp introduced his presentation by pointing out that there is still much to discover regarding molecular systems and the role of the RNAs. This is particularly true of microRNAs (miRNAs), which regulate protein expression by targeting specific mRNAs. An understanding of the full transcriptome will be important in order to answer questions related to cell states and tissue-specific protein expression in normal and tumor cells. The next generation of massively parallel sequencing techniques will soon make this evaluation economically as well as technically feasible.[10]

The importance of miRNA regulation of biological pathways. Bioinformatics studies have found 250 to 1,000 genes that encode miRNAs; these miRNAs probably regulate up to 50% of all genes in vertebrates.[11-15] miRNAs regulate biological pathways by binding mRNA and regulating translation; 25%-50% of all mRNAs interact with miRNAs. mRNA sequences enriched in complementary sequences to miRNAs have also been found, and there are 87 evolutionarily conserved seed families of miRNAs. The distribution of preferentially conserved target sites in the 3′ UTR includes 55% of genes with one or more target sites, while 45% of genes do not have sites.

Studies using DNA expressed sequence tags have demonstrated several methods for variation of transcripts coming from a single locus. These include standard transcriptional activation, alternative promoter usage, exon inclusion/exclusion, and 3′ UTR utilization. Using high-throughput sequencing data, approximately 94% of human genes, or essentially all multiexon genes, are estimated to undergo alternative splicing. Of these, more than 90% undergo alternative splicing with a minor isoform fraction of at least 15%. More specifically, there is evidence for tissue-specific regulation of splicing. Of the eight common types of alternative splicing that make up 70% of regulated expression, sequence conservation is associated with switch-like exon expression. Context- and tissue-specific activity can be inferred from the patterns of motif conservation flanking tissue-regulated exons.



The role of miRNA in cancer. Loss of miRNA regulation has been correlated with cancer progression. Furthermore, changes in specific miRNA molecules have been identified in cancers. For example, miR-15a and miR-16 downregulation is seen in Stages 2 and 3 prostate cancers. These miRNAS are found at chromosomal region 13q14, which is frequently deleted in cancer. Their tumor suppressor role is suggested by observing that their expression in primary prostate cancer cell cultures is inversely correlated with expression of proteins associated with cell survival (BCL2), proliferation (CCND1), and invasion (WNT3a).[16]

In summary, a large percentage of human genes undergo alternative splicing, a majority of which are tissue regulated with a substantial amount of individual-specific variation, leading to the question: Is this a mechanism to coordinate the open reading frame (ORFeome) with the untranslated regions (UTRome)? The switch-like exons have distinct protein coding and conservation properties, suggesting important functions. There is also evidence for coregulation of splicing and cleavage/polyA events.

Group Discussion: Cancer Information

Christoph C. Adami, Ph.D., David Haussler, Ph.D., M.S., Phillip A. Sharp, Ph.D., and Group

With reference to the presentations of Drs. Adami, Haussler, and Sharp that focused on molecular information, the group explored the context in which this information should be used to study cancer. Key concepts raised were as follows:

Are we missing the forest for the trees in looking at all small changes? Is this approach useful for developing treatments? How does knowledge of small changes inform development of new interventions for cancer treatment? It was argued by some participants that this level of granularity is needed to get useful treatments. We also need to consider other molecular factors, such as epigenetic and posttranslational modifications, and incorporate new approaches to integrate the huge amounts of data coming from new high-throughput analysis systems. It was pointed out that evolution finds modular solutions (e.g., multiple pathways); the study of these modules would yield some insights. Pathway analyses may offer new approaches to developing better therapies. As tumor cells are selected, they become information rich and pathway dependent; identifying these pathway dependencies is important. To incorporate information theory into this line of research, stochastic modeling of pathways could be developed—perturb them and measure outcomes. We also need to incorporate time into the analysis.

The discussion turned to the potential "normalization" of cancer cells, viewed as an interesting and challenging concept. If we knew how to regulate transcription factors (and it was agreed that we do not have this knowledge yet), a cell could in theory be normalized. The major factor to consider in attacking cancer from this standpoint is determining how much information comes externally from the niche and how much internally from the cell.

The group discussed the concept that evolution in cancer is different from evolution of a species. The exciting difference is that species evolution occurs over thousands of years and cannot be "redone" in order to study it. Conversely, we can watch cancer evolution in the body. We can repeatedly observe how changes happen and see the same changes over and over again. For example, all the p53 mutations (adaptive mutations) are repeated adaptations of a similar type. This is analogous to convergent evolution. In this context, it was pointed out that age is the most carcinogenic event, and perhaps system decay could be an interesting model in which to study cancer development.

Small Group Discussions:

Information Theory—If It's So Important in Cancer, Why Have We Not Made More Progress in the Field? (Robert Mittman and Group)

For a full graphical representation of this discussion, see Figure 6, Appendix 1.

In this first brainstorming session, the participants were invited to work within small interdisciplinary groups to identify research questions that might be addressed using the concepts discussed: the nature of biological information, information flow, translation, and information theory. The groups generated a large number of research questions, summarized as follows:

- Are there patterns in sequence and expression data that represent the cancer state?
- What is the number of meaningful states in cells? What is the meaningful level at which to characterize them?
- What are the information channels necessary for cancer progression (does cancer proliferation occur through channels)? Where is information stored in cells and tissues, and what is the relevant time information about cancer progression? What is the important vs. unimportant information to gather at levels?
- How can we use information theory to diagnose or predict cancer? How do we extend information theory to encompass survivability (fitness) of tumor vs. normal cells? Is cancer an increase or decrease of information or entropy?
- How do we incorporate function into information theory? How does one develop a precise notion of context (cell niche)? Is there an information characterization for "stem-like cells"?
- How do we design meaningful experimental model systems to capture interactions between tumor and normal cells?
- What are the right tools to measure specificity and sensitivity of cells to a time-dependent environment? Is it possible to phenotype a cancer through distal molecular measurements? Can we use computer software and hardware verification methods to probe cells?
- How do we integrate information theory with precise measurements?

Panel Discussion (Brief Presentations)

Contextual Translation of Information: So Many Signals, So Many Channels, So Much Translation on So Many Scales

For a full graphical representation of this discussion, see Figure 7, Appendix 1.

In this session, participants heard four short presentations representing different perspectives on the uses of information theory in studying cancer. These presentations moved the focus of the discussion from the molecule to the larger scales of organelle, cell, tissue, and organism.

Beyond the Genome: Understanding the Human Somatic Cell Tree, Somatic Cell Molecular Clocks, or "Hey Doc, How Did I Get My Tumor?"

Darryl K. Shibata, M.D., Professor, University of Southern California

Dr. Shibata started the panel by describing how the information translation at the molecular level during somatic division, the fidelity in the epigenetic DNA methylation replication patterns, or noise can be used as molecular clocks.

DNA can be viewed as an information molecule containing a set of instructions and historical information. If the information source is the zygote, stem cells are both the transmitters and receivers, and the current cells are the destination. As transmitters, the stem cells make copies—daughter stem cells—as well as differentiated cells that eventually die. At the stage where the new cells become either stem cells or differentiated cells, replication error, or noise, can occur. The "molecular clock hypothesis" is that cell copies contain replication errors proportional to their mitotic age.

As a tumor becomes more diverse, the analysis of cells across the tumor should provide more knowledge about the history of the tumor; diversity = antiquity. Epigenetic methylation clocks or measurement of age-related increases in CpG DNA methylation can be used to understand the history of the tumor. Counting the difference in cells on different sides of a tumor reflects the number of replications and thus the age of the tumor. While methylation is removed early in development in some tissue, age-related increases in DNA methylation occur in mitotic human tissue, such as the colon, and can be polymorphic. Thus, methylation pattern diversity may represent replication errors of drift.

Analysis of methylation patterns may represent an approach to test the hypothesis that chemotherapy failure is due to preexisting resistant cells. Dr. Shibata also raised the possibility that a younger cancer might be less diverse and more responsive to chemotherapy and an older cancer more diverse and less responsive.

Finally, he stated that somatic cell histories are likely recorded by replication errors in their genomes. It should be possible to translate modern molecular phylogeny approaches to somatic cell "evolution." While many practical problems remain with implementation, the approach would be clinically useful.

Signaling Pathways: An Engineer's Perspective

Philip R. LeDuc, Ph.D., Associate Professor, Carnegie Mellon University

Moving up to the cellular level, **Dr. LeDuc** spoke on the usefulness of modeling translation of input and output signaling information at the cellular level, which he studies from a mechanical engineer's perspective, as systems. He discussed the value of building models to understand biological systems, pointing out the similarities and differences between cells and robotic systems.

The cell processes environmental cues from a wide variety and large number of inputs and uses control and feedback loops to produce outputs (such as apoptosis, motility, quiescence, etc.). Furthermore, cell processing and signaling involve spatial and time dynamics and a huge number of molecules. The robustness of the system is a key factor, as well as signal integration and noise. Noise in biological systems can stabilize a system in the context of many incoming signals.

Thus, it is important in modeling cancer cells to define inputs and outputs. Important related areas include feedback, feed-forward, and integral control. Dr. LeDuc's experimental approach is to build spatiotemporal control into models with the use of microfluidics. He also is interested in modeling cell crowding and tissue implants.

Multiscale Nature of Information Transfer

Mauro Ferrari, Ph.D., M.S., Professor, University of Texas Health Science Center at Houston

At the patient level, **Dr. Ferrari** discussed use and translation of the information imbedded in the biological properties of the body's transportation systems to optimize a new generation of drug therapies. Dr. Ferrari's interest is in information transfer in biological systems from the health care perspective, in particular, information transfer from the physician to the cancer and back.

The recent information revolution in the human world has been triggered through communication at the chip (electronic) level, which is spatially directed, with built-in time sequences. Conversely, recent work in biological systems demonstrates that communication within biological systems (cells, organelles, etc.) is not as spatially directed but is based on biological specificity. This is also applicable to the communication between the physician and the patient's cancer and how the information is managed. For example, if one starts with an injection of a drug somewhere in the body, the drug has to somehow travel to the target (injection to location, point A to point B). The term "drug delivery" is oversimplified. Although the drug may have high specificity for the target, the transport process is complex, containing information in the steps between point A and point B. For example, the transport may include avoidance of undesirable uptake, metabolizing, and clearance mechanisms, as well as navigation through normal circulatory pathways and tumor vasculature. Thus, the drug transport pathway uses many forms of communication involving biophysical transport (active transport, diffusion) across biological barriers. These transport modalities are part of a transportation code. The next stage in development of drug therapies will make use of biological properties/transportation systems to optimize specificity and transport modalities (e.g., P-glycoprotein-mediated transport). Thus, the sequence of code used to manage transport through biological systems across biological barriers is of significant interest for future research.

In addition, from the perspective of the physician/cancer communication pathway, it is clear from use of tools like ultrasound that signature differences across normal tissue and cancers are architectural. Improved diagnostics are needed to explore the different architectural signatures with drug response. In particular, development of 3D multiscale (macroscopic and molecular) mathematical modeling tools generating models that are consistent across these scales would aid cancer treatment investigations.

Dynamics and Crosstalk of Intracellular Organelles

Jennifer Lippincott-Schwartz, Ph.D., M.S., Senior Investigator, National Institute of Child Health and Human Development

At the subcellular, organelle level, **Dr. Lippincott-Schwartz** discussed implications for cancer in mechanisms of nongenomic cell cycle regulation.

Dr. Lippincott-Schwartz described her studies on mitochondrial regulation of cell cycle control, including p53 involvement and the implications for cancer therapy. Her work also illustrates the value of examining not only the genomic code but also the use of traditional cell biological approaches to understanding the role of nongenomic cellular processes and cell organization in cancer.

She has experimentally demonstrated that mitochondria change morphology with cell cycle. The organelles take on a hyperfused morphology at G1–S (similar to Dynan mutants that prevent fission). Additional properties are an increased matrix continuity, electrical connectivity, and maximal adenosine triphosphatase production vs. other times in the cell cycle. Depolarization prevents mitochondria from reaching the hyperfused state, which results in preventing cells from going into S-phase. This is the only time in the cycle that cells are sensitive to mitochondrial depolarization. There is also evidence that the fused mitochondrial state leads to a buildup of cyclin E, without other cyclins accumulating. p53 may play at least two different roles in the system. If the mitochondria are depolarized, a p53/p21 block occurs, but p53 may also independently control genes involved in mitochondrial respiration.

Mitochondria, with p53, may regulate a restriction point in the cell cycle at G1–S progression; investigation of this may be an opportunity for cancer therapy.

Information Theory in Living Systems: Contributions of the Microenvironment

Robert Gatenby, M.D., Division Chief, Moffitt Cancer Center and Research Institute

At the cellular level, **Dr. Gatenby** discussed the dynamics and continuing optimization of information flow in nongenomic structures and the revaluation of that information in the cancerous state.

Dr. Gatenby described the application of information theory to understanding cellular stores of information and their relationships to cancer. He noted that information theory is limited by **context**; thus, **key issues are the value (or fitness), reception, and cost of the information**. In terms of **context and reception**, biological information requires both order and meaning as the flow of information is from a sender to a receiver. As such, optimization dynamics are continuously occurring, maintaining only enough information for the cell to function.

In the study of information in cancer, there is a balance of context and cost. For example, seemingly similar kinds of information may have more value than others. Differentiated functions of cells (high information and energy) come at a high cost but create high value in terms of maintaining the viability of the organism. However, for the transformed cell, a differentiated function has high cost but low value in that it does not contribute to cellular proliferation. Therefore, cancer cells will tend to lose differentiated functions but gain information that promotes proliferation of the individual cell.

The integration of thermodynamic buffering of the cell into control mechanisms is important. How a cell maintains constant entropy may be due to the varied mix of information that cells maintain. One component of cellular information is the DNA-RNA protein system that encodes heritable information. In addition, critical and important information may be encoded in some of the nongenomic centers, such as membrane content, membrane gradients, all highly nonrandom structures, and cytoplasmic information sources. Cells maintain an ensemble of integrated information units that constantly assess the state of the cell, including regional and temporal environmental and cytosolic functions.

Thus, in cancer, the fundamental dynamics are flow of information into and out of the cancer cell; critical information may be encoded in nongenomic structures of the cell. There is then a continuous optimization process of the cost and fitness benefit of each information bit. Carcinogenesis is fundamentally a process in which information is revalued.

Discussion Highlights: The panelists and other participants discussed several questions posed to stimulate thinking on cross-scale application of information theory to understanding cancer. Following are questions and some of the highlights that emerged from this session:

- **Can cancer be reversed and/or the cells "normalized"?** Cells could be shifted back if control of the cell cycle could be regained (e.g., as suggested by Dr. Lippincott-Schwartz's work on mitochondria). In vitro experiments have also sought to reverse the neoplastic process by placing tumor cells in nontumor environments. All of these studies suggest that the context in which the cell exists is important.
- **How does one define information flow, and how can information theory be employed to understand the intercellular signaling that exists in microenvironment (including stroma cells, etc.)?** Tumors require maximal information and unique flows of information; communication between cells is critical for cells to proliferate.
- **In tumors of different types, some cells are full of mitochondria, so do we need to know how cells control mitochondrial proliferation?** Cellular synthesis of precursors was suggested as one point of control and a logical area of investigation.

Small Group Discussions:
Understanding Signaling and Contextual Translation of Information at Multiscales: What's Relevant From the Physical Sciences?

For a full graphical representation of this discussion, see Figure 8, Appendix 1.

In this second brainstorming session, small interdisciplinary groups were asked to consider information received from prior sessions along with potential physical mechanisms to revisit the earlier question of the most relevant research questions to unravel the complex information associated with coding, decoding, transfer, and translation in cancer. The output from the groups increased to include other questions as follows:

- What are Shannon's channel and noise in terms of DNA and its actions?
- How much inheritable information is encoded solely in the intracellular structures of normal and cancer cells?
- To what extent do the miRNAs have paracrine signaling functions, and what are their roles across scale (DNA/protein/cell)?
- What are the mechanisms by which the signals are modified out of the cancer cell? What is the minimal set of information required for cell-cell and cell-matrix communication in cancer?
- What tools do we need to predict and control multiscale communications? Can we develop tools to measure spatial and temporal variation and intercellular and intracellular gradients?
- How can we make predictions about increasingly complex cell behaviors and build increasingly complex models of cell behavior to understand cancer (e.g., such as for cell movement)? What are all the factors that affect the cell cycle in cancer?
- How is cancer initiated: at the cell level (single abnormal cell) or by a change in niche at tissue level? What are the phylogeny and phenotypes of premetastatic-metastatic tumors?
- How do we characterize tissue niches (elasticity, etc.)? How do we measure the physical forces that define these niches?
- Precisely how is the information energy burden in cancer calculated?
- Considering the tumor as an ecosystem, there are key dependencies between cells. Do these differ in low-grade, well-differentiated tumors vs. high-grade tumors with poor differentiation? How do we understand interdependencies of tumors at the tissue level? If we can analytically define cells (gene expression, tissue organization in organisms such as *C. elegans*, can it be done for a mouse tumor?
- Why do patients with cancer die? Is this a specific aspect of information precipitated by or controlled by cancer cells? Are there systematic ways to extract predictions from higher level (tissues, organs) descriptions of cancer?

Panel Discussion

The Outcomes and Consequences of Information Transfer in Cancer Across Length Scales

For a full graphical representation of this discussion, see Figure 9, Appendix 1.

The group next heard four short presentations directed to the transfer of information across the various scales.

How Information Is Used To Build Cells: Design Principles and Information Transfer

Wallace F. Marshall, Ph.D., Assistant Professor, University of California, San Francisco

Given that cells and organelles are extremely complex, Dr. Marshall raised the question of whether the genome has all of the information needed to specify this degree of complexity. He posited that the genome may not be a blueprint for cellular structure, given that blueprints are geometrically explicit, position-based plans, without timing or order information for the building process. Conversely, **the genome is timing based, with the geometry implicitly based in the genome**. Thus, the question arises as to how much of the genome is needed to specify construction of the cells. By using a model system to explore the determinants of organelle structure, **Dr. Marshall** argues that cells are probably not as complex as they appear; a limited number of genes may determine the complexity of cellular structures.

In order to probe how much information is needed to build structures at different size scales, specific case studies of subcellular structures can be used to identify design principles that underlie cellular architecture and assembly. The approach is to use a **simplified organelle-level description of cellular structure**. This approach avoids attempting to work at the detailed and complicated level of biochemical pathways, facilitating the study of size, shape, number, position, and orientation of the organelle. If structural information needed at the organelle level is understood, theoretically it could be put together to obtain understanding of the overall structural information requirements of cells and how much of the genome would be required to specify cell structures.

As an example, Dr. Marshall used his studies of the dynamic maintenance of flagellar length in cilia to examine one structural component—size. Cilia are microtubule-based structures; length is maintained at the correct rate by a steady-state process that assembles and disassembles subunits using intraflagellar transport (IFT) rafts. Since cilia are linear organelles, cilia length makes a good model for studying organelle size; it is relatively less complex than size in other organelles (which can be dependent on volume, structures, etc.). One question is what information is needed to achieve and maintain a defined cilia length? The goal of the control system, assuming that cilia are at the correct length, is to have equivalent rates of assembly and disassembly; perturbation of either would change the equilibrium steady state and length. Dr. Marshall employed genetics to study length control, using multiple mutations to demonstrate that a single component, assembly rate, determines the length. The disassembly rate is length independent, while the assembly rate is under a control mechanism and limited by IFT, which is inherently length dependent. Two mechanisms result in a longer flagella mutant phenotype, increasing the assembly or decreasing the disassembly rates. Multiple mutations can produce the same result. This work demonstrates that multiple "calculations" in a cell can lead to a single end point.

This study illustrates that complex cellular structures can be effectively studied using simple models. Moreover, Dr. Marshall argued that evolution would favor the use of crude schemes using fewer rather than more components and that cells are less complicated than they appear, in that most organelles

probably require a small number of genes to modulate their geometry. It remains to be seen how this model will be applicable to organelles in general and to cancer specifically.

Intersection of Evolution and Information Theory: What Does It Mean for Cancer?

Carlo C. Maley, Ph.D., Assistant Professor, The Wistar Institute

Evolution can be thought of as an algorithm for creating and transferring information. Mutations generate new variants, and natural selection eliminates the maladaptive variants, leaving a correlation between the genome and the environments in which it evolved. Cancer is one example of multilevel selection; the tumor suppression mechanisms generated by billions of years of evolution can be dismantled by somatic evolution within a human lifetime.

Three factors are considered necessary and sufficient for natural selection. These factors are observed in all clones in neoplasms that have a phenotype that is favored over other phenotypes. Dr. Maley outlined these factors as follows:

- **Variation in cell populations from somatic mutations.** Genetic heterogeneity within neoplasms is commonly found and well documented; somatic evolution can give rise to heterogeneity.
- **Heritable variation among cells.** Encoded genetic and epigenetic changes are carried over to daughter cells during cell division. Clonal expansions are the signature of expansion of neoplasms and can predict progression.
- **Variation that affects fitness, reproduction, and/or survival of the cells** (e.g., suppression of apoptosis). It is also important to note that the **fitness effect of the mutations is also a function of the microenvironment**.

With regard to somatic evolution, human cells are well adapted to being part of the cooperative environment of a multicellular body, but they are not initially well adapted to being a cancerous parasite within the body. Thus, the starting point for a cancer is likely far from the optimal point for the cancer. However, since most mutations would probably be deleterious, it is not clear that that is true for the neoplastic cell. Some mutations must affect the genes responsible for differentiation and cooperation. Dr. Maley questioned what percentage of mutations increase the fitness (survival) of a somatic cell.

Although the evolutionary view of cancer has been around for decades, the field has not developed as needed. As a result, many questions about details in the evolution of neoplasms remain unanswered, and a significant amount of work is still to be done. Questions involve mutation rate, population size, and generation time of cancer cells. How long does progression take on a single cell/tissue basis? How much population structure is in a neoplasm? What are the selective effects of mutations? How does the microenvironment change those selective effects? What are the selective effects of our therapies? How does the configuration of clones change over time?

Relative to information theory, it is interesting to note both information gain and loss in cancer. Most cancers not only have extra DNA (are hyperdiploid) but also have large losses of genetic information and large regions of homozygosity. This suggests that reversibility of cancer cells is questionable, as there is no way to gain back the information.

Furthermore, Shannon's information can be measured within a neoplasm by characterizing the number and frequency of clones; Shannon's diversity predicts progression. For example, measurement of the frequency of clones in a Barrett's esophagus neoplasm found that neoplasms containing more variability (Shannon's information) were more likely to progress to cancer.

In conclusion, information theory and cancer are connected, since the transfer of information over time occurs during neoplastic progression or via evolution. Evolution builds information only in heritable structures; heritable changes in neoplasms include genetic and epigenetic changes. There are many forms of information in cancer (e.g., signal transduction from the microenvironment). Information is both created and destroyed by somatic evolution in neoplasms, and this process drives neoplastic progression and accounts for therapeutic resistance. That process is poorly understood and represents a huge opportunity for progress in cancer research.

The Physics of Information Transfer in Cancer

Robert H. Austin, Ph.D., Professor of Physics, Princeton University

To study cellular interactions and information transfer involved in cellular survival, **Dr. Austin** employs the study of bacteria in complex microenvironments using nanofluidics and arrays.

Bacterial mutants that evolve in environments with unchanged culture media adapt to stress and express a growth advantage in stationary phase (GASP), emerging as GASP mutants. When the culture medium lacks nutrients and the two types of bacteria are "stirred" together, the wild-type bacteria reduce metabolism and conserve resources, while the GASP mutants do not decrease metabolism and will overgrow the wild type, similar to a cancer. However, depending on how the two strains are mixed, a complicated interdependent relationship is observed. Both cell types can grow well together, due in part to mutual benefits obtained from proximity; for example, the mutants are able to metabolize wild-type waste products. Cell clustering of the strains can be analyzed using the Pearson Correlation Coefficient (1 indicates attraction, –1 repulsion, 0 chaos). In culture, the GASP strain forms relatively diffuse clusters, while the wild type develops tight clusters with correlation coefficients near 1. If the two strains are mixed, the coefficient changes over time to –1. **Thus, when the two strains are mixed, fitness for both forms is optimized by clustering through nonself-avoidance and self-recognition and communication; crosstalk between species becomes apparent, and fitness is optimized by coexistence at different length scales.** The length scale of the interaction between GASP and wild types not only is local but also reaches metascale correlations.

These experiments also demonstrate that **both strains are necessary for the stable existence of the species in the presence of the complex environment**. This leads to the question: Is cancer a necessary defense mechanism for the species? Dr. Austin suggested that information approaches and theoretical constructs may help explain the language of coexistence and cooperation in more complex systems.

Information Theory: Could This Approach Enable an Understanding of the Why/How of the Malignant Phenotype?

Christoph C. Adami, Ph.D., Professor, California Institute of Technology

Dr. Adami discussed how cancer research could take advantage of the context dependence of information in cancer through examining interactions of genes in cancer pathways to evaluate critical mutations in cancer. Because somatic mutation rates in cancer are often elevated, not only are oncogenes and tumor suppressor genes mutated, but also other less significant genes that may not impact cell transformation are also mutated. Therefore, the issue is how critical mutations that cause cancer are distinguished from mutations that are just associated with cancer. Because the fitness of one gene can be contingent on the fitness of another, the same method to find important protein channels described in Dr. Adami's keynote address above can be used to find important mutations in cancer—look for the signals that change. In theory, to define the channels between proteins in cells, all the protein combinations would be tested. However, a more practical approach is to look for changes in signals from proteins that are actively signaling. Critical mutations in cancer can be investigated in the same way, and Dr. Adami proposes that genes that interact in cancer can be viewed as a network.

Given two genes, an oncogene and a tumor suppressor gene, each has wild-type information content and wild-type replication rate. If information can be used as a proxy for fitness, then a mutated oncogene, a faster replicator, should have a mutation that has increased its information content. In addition, the oncogene increase is conditional on another mutation in a gene within a pathway deactivating a tumor suppressor pathway. If the mutation is within a protein, that protein's information content may be decreased. Thus, **the increased information of the oncogene is conditional on a mutation deactivating a tumor suppressor pathway, and so only the correlated mutations between oncogenes and tumor suppressor genes are diagnostic of cancer**.

Dr. Adami proposed that finding a network of genes that interact to change replicatory fitness is tantamount to discovering cancer pathways. The cancer pathway requires finding correlated mutations or linked gene products within a pathway. Correlated mutations can be identified using information theory if linkage information exists, such as whether two genes with mutation patterns are present in the same cell. Similar investigations with HIV protease and reverse transcriptase demonstrated that correlated mutations happen only in sequence regions with high entropy and are not more likely to happen than by chance. However, Dr. Adami predicts that correlated mutations linked between proteins in a cancer pathway are more likely to happen than by chance, because they are associated with a cancer genome.

The current sequencing paradigm focuses on identifying genes that have a significant number of mutations. Of note, application of the program requires new guidelines on data collected per patient. Patient-specific lists of mutations and profiles are needed for such correlations. Finally, information theory can also be used to track and study drug resistance mutations in cancer, just as for the example of drug resistance in HIV described previously at this meeting.

Group Discussion

Robert Austin, Ph.D., Christoph C. Adami, Ph.D.

As in earlier discussions at this meeting, the panel members considered several key questions related to the panel's topic. Key points and some points from the discussion addressed included:

Although many cells circulate per day that are sloughed off from primary tumors, very limited numbers of these cells result in metastasis—why is that true? Some points from the discussion are summarized below:

- There is currently no known mechanism by which a cancer cell can transmit a cancerous phenotype or imprint onto a preexisting normal cell. Therefore, cancer arising in remote locations can be understood as derived from cells sloughed from the tumor that are derived from stem cells. These sloughed-off tumor cells may become nonmalignant for a period of time but can eventually transform back to be like stem cells.
- One explanation may be that the microenvironment affects phenotype and favors metastases. A preconditioned niche may facilitate neoplastic growth, but this phenomenon is not well understood. There are certain spots where tumors settle, consistent with preconditioning.
- Infectious processes are associated with cancer (e.g., *Helicobacter pylori*, which has been related to inflammatory processes). Also, women with breast cancer given bone-promoting drugs after their breasts are removed have their risk of recurrence reduced by 50%, similar to the risk reduction following 6 months of chemotherapy. Both these observations are consistent with indirect effects (i.e., changes in microenvironment) as critical factors in determining the growth of cancers.

Cancer tumor growth involves behavior similar to that of punctuated evolution. What causes punctuated evolution?

- Punctuated evolution is used to explain the fossil record. It is not well understood as applied to organisms, much less cancer, but a current theory is that populations achieve a fitness plateau; neutral mutations produce a similar phenotype and fitness and then pop to a new plateau.

Are there emergent properties in evolution and cancer?

- Surely, it is believed that nonlinear interactions occur in all complex systems.

Are a critical number of cancer cells needed for metastases to occur, and can this be modeled using concepts from phase transitions?

- One piece of evidence is the finding of minimal residual disease in leukemia patients. Observations show that after therapy, cancer cells can still be detected in the blood (BCR-ABL), but the patients remain stable for a number of years. It is still not known whether, if leukemia remains below some critical cell number, the immune system is able to control it. If so, an area of investigation would be to understand the cooperation dynamics that lead to density effects.

Are cancer stem cells reality or fiction? Do stem cells lead to cancer, or do cancer cells behave like stem cells? Are cells in a tumor heterogeneous? How would the stem cell concept be reconciled with cells in a tumor having the same genome?

- Cancer stem cells are a powerful idea.
- The concept of stem cells is orthogonal to the evolutionary approach; no doubt there are different phenotypes in a tumor, so what is the population of evolving cells? If stem cells exist, that means the evolving cells have been reduced to the stem cells only.
- Is the question one of frequency of stem cells, as stem cells are cells that proliferate for a long time?
- Is it also possible that the information needed for cells to become neoplastic requires changes in cooperation in a network (cooperative activity/information/mutations) rather than getting enough mutations to escape control?
- It is hard to see how selective forces would operate in an environment that does not yet exist. Although niche signals (growth signals) are mysterious, it is clear that they are ubiquitous and already readable by cells. Thus, a mutation to use these signals in cells as a metastatic mechanism would be consistent with cooperation.
- The fitness effect of one mutation may be dependent on a number of others happening first (i.e., it may be a combined effect of several "neutral" mutations).

Panel Discussion

The Future: If We Understand the Specifics (Physics, Chemistry, etc.) of the Information, Its Transfer, and Contextual Translation at Multiple Length Scales in Cancer, Can We Alter Outcomes?

For a graphical representation of this discussion, see Figure 10, Appendix 1.

Paul Davies, Ph.D., D.Sc., Professor, Arizona State University; Donald S. Coffey, Ph.D., Professor, Johns Hopkins University; Robert Phillips, Ph.D., Professor, California Institute of Technology; W. Daniel Hillis, Ph.D., Chairman, Applied Minds, Inc.; John E. Niederhuber, M.D., Director, National Cancer Institute

In this session, each panel member was asked to pose critical questions to other members of the panel. The general subject was using information at all levels to affect outcomes. Following are the questions posed and highlights of the ensuing discussion.

Will we have enough information and processing power to manage cancer without ever really understanding the problem? In other words, will we ever understand the complexity of cancer? The panel's discussion is summarized below.

- We need some principles to understand how to apply computational methods to solve the problem. It is possible that we will never completely understand the problem, but having no solution is not an answer. In certain systems, achieving some level of understanding will enable achieving computational control over a system. In addition, we will need both theoretical constructs and computing power to achieve this level of understanding (whatever it may be).
- It is worthwhile comparing tumors we can successfully treat with those that are currently untreatable. Empirical approaches are useful; for example, information transfer develops the fertilized egg into a chicken in the presence of heat. This transition requires time and energy and is dynamic; once developed, the chicken must sustain energy levels to live. Phenotypes, including the cancer phenotype, can be reversed with heat. The basis for all cancers is a morphological transition, demonstrated by introducing heat-sensitive SRC mutants into cells. By changing the temperature, the cells can be forced to change between normal and a tumor-forming, cancerous phenotype. This is pertinent, since all protein, DNA, and RNA folding is temperature sensitive. Temperature regulation may be one approach to address this complex issue from a different angle. Thus, the big question is how heat regulates information in a cell.

Given the above, is it possible to develop simple phenomenological models of cells, with a few (10-15) parameters to fit in order to define cancer?

- A lot of progress can come from the use of simple models, but fitting a large number of parameters may be difficult. Fifteen seems to be too many.
- Although difficult, more parameters may be needed to model cells to account for tissue-specific complexity. This would be possible, because the only way complex systems have been controlled has been with very simple models.
- In evolution, extinction is driven by changing the habitat, not by random mutation. We should focus on changing the habitat and not the tumor cells, because tumor cells become resistant to drugs, while normal cells do not.

What will it take to understand the heterogeneity of the tumor? What if most of the information in the tumor is noise and only a small subset of the cells is crucial for carcinogenesis, meaning that a lot of the aberrations are not important? Perhaps cells that are more organized are better able to renew.

- Cell vibrations are observed in cancer, and it has been suggested that stochastic resonance is involved. What is the importance of stochastic resonance in cell signaling? Does the noise level need to be raised to see small signal peaks? The evolution of the nucleus initially involved primitive keratins and laminins, and these molecules continue to be implicated in cell structure and signaling. This suggests that a lot of cell structure/organization is designed and aligned for cell signaling. Can a cell tune itself?
- Perhaps we should focus on similarities between cancer cells, not differences. The question of noise underscores that the deepest levels of detail are often not useful in biological systems. Modeling at a level between minimal and deepest complexity is likely best for control.

Is the notion of a cure for cancer a meaningful concept, and would a Manhattan Project for cancer be viable?

- We do not have to cure cancer, just manage it. We can cure some cancers today, but it is doubtful that we will eliminate cancer. However, we will make progress in control and prevention. The progress in control will come with understanding the system and microenvironment in which the tumor exists. We will learn about the roles of tissue progenitor, stem cell, and viral infection. Currently ~20% of cancers are known to have viral involvement, a percentage that will likely increase.
- Understanding cancer evolution may be most helpful. All organisms have DNA, but in extreme environments (e.g., in the deep sea, mines under high pressure, radiation, and temperatures), some organisms survive because they have evolved systems that protect against stress. We need to understand how evolution works to allow organisms to survive and how stress systems work.

In relation to this meeting, is the interest in understanding cancer or in successfully controlling it? Are these interests tied together, or in fact quite different?

- It was essentially agreed that it was not necessary to understand every aspect of cancer development in order to intervene successfully. A short brainstorming discussion ensued regarding the models that might be used to develop understanding of various aspects of cancer evolution, information, and complexity. It was suggested that a model capturing regulation of proliferation would be valuable. Regarding evolution, it was noted that mathematical modeling done to date describes what has happened in the past, but we require models that predict future events. A suggestion was made that cancer might be viewed as a quasi-species in terms of evolution and information. It was noted that one of the key forces in evolution is development of modularity, which should be examined in cancer. Other interesting questions for investigation are multicellularity and hierarchy. Development of a model of phase transition in cancer may be helpful.
- Dr. Austin's experimental system with bacteria (described above) has potential for use in evolution studies (i.e., to look at changes over time). It would also be good to convert the system to use with somatic cells.

Mr. Mittman closed this session by asking the panel members to comment on what they would like to see addressed in the rest of the meeting. They responded as follows:

- Discuss cell communication and heterogeneity and how to quantify and model these events mathematically.
- Detail specific theories to test.
- Sharpen definitions of information in cancer.
- Address cancer control systems.

Day 3: Friday, October 31, 2008

Meeting Review and Introductions

Anna D. Barker, Ph.D., Deputy Director, NCI

Dr. Barker noted that the meeting had been interesting and challenging given the complex and broad subject matter of information in cancer in the whole arc of coding, decoding, transfer, and translation, in addition to the focus on communication among the various length scales and across time. NCI is funding a large number of scientists to research the various aspects of contextual information in cancer; however, she and Dr. Niederhuber are now trying to bring a missing piece into this research, the physics of the process.

In meetings such as this, a conceptual framework should be built around what has been discussed. The overarching question of whether information theory has a role in understanding cancer, posed at the start of this meeting, will be revisited. In particular, during the presentation by Dr. Hillis, Dr. Barker asked the group to think about some of the following questions that had been raised throughout the meeting:

- What is information in cancer? We need to start to answer this question. Think about the question, What is a gene? It is probably not what we thought it was.
- How does information in cancer cells differ from information in normal cells (if at all)?
- How do cells transfer information (purview of the physicists)?
- How does one actually interpret information at all the length scales?
- If we knew the answers to the questions above, would information theory make sense as an organizing principle? (vs. other possibilities, e.g., algorithmic solutions?)
- If we knew some of these answers, would it change the way we diagnose, treat, and prevent cancer?

Dr. Barker then introduced **Dr. W. Daniel Hillis** to give the final keynote presentation of the meeting.

Keynote Presentation

The Failure and Repair of Emergent Systems: A Systems Engineering Approach to Cancer

W. Daniel Hillis, Ph.D., Chairman, Applied Minds, Inc.

Presentation Highlights (For a full graphical representation of this talk, see Figure 11, Appendix 1.)

- Emergent systems are composed of subsystems at multiple scales and complexity, working together to produce the emergent properties of the whole.
- Emergent systems are incrementally created, are local and repetitive, and have robustness and order at multiple scales.
- There are three control systems in cancer: the patient's body, the cancer, and the patient treatment loop.
- Direct treatments to help the body win over the cancer:
  - Optimize information bandwidth in patient-physician communication channels; information measurements are key.
  - Eliminate extraneous levels of meaning; treat, do not diagnose, disease; look for clues for treatment choices instead of biomarkers of disease.
- What is understood is not necessarily the best level to control.
  - High information effectors do not always correspond to understandable patterns.
  - Treatments with high degrees of freedom may be the most effective.

Dr. Hillis started his presentation by describing emergent systems in general to set the stage for an illustration of how this approach can be used to devise an alternative approach to the treatment of cancer patients. While he acknowledged that his suggestions might be viewed as naive or heretical by the oncologist, the goal of the approach is to control the cancer. From his perspective, understanding the details of cancer biology, while valuable as a tool, may not be required to reach this goal.

Emergent systems, such as computer networks, organisms, and economies, are complicated. As such, systems as a whole have behaviors that are not obviously deducible from behavior of the component parts. A system can be viewed as a black box with inputs and outputs and states. Analogously, a cancer patient can be viewed as a system with various inputs (diet, treatments) and outputs (indications of health) and a goal state (health). Most of the state is hidden, although clues can be measured, such as a patient's temperature. In addition, the state transition functions are also unknown and hidden until the mechanisms of the organism and cancer are understood.

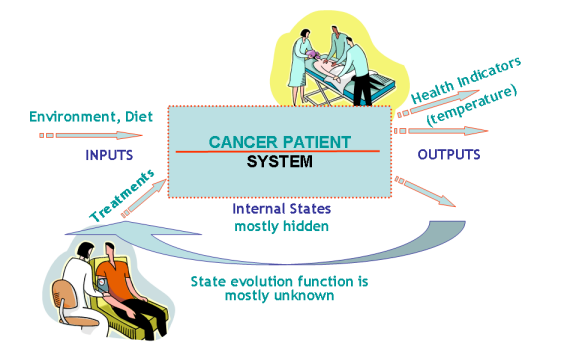
Furthermore, these systems can contain subsystems at multiple scales and complexity (molecular, cellular, tissue, organism levels). The components work together to produce emergent properties of the whole that are not reflective of the properties of the parts. An example is life, in that biomolecular interactions in cells lead to the property of life in an organism. In other words, emergent properties are the things we care about but do not tend to understand at a mechanistic level.

Emergent systems have common properties that can be studied, including:

- **Incrementally created:** For example, by processes like evolution or design.
- **Locality:** Parts tend to interact with only a few other parts in a meaningful way.
- **Repetitive:** Parts tend to have the same themes, but with variations such as subvariables.
- **Order at multiple scales:** Molecule, cell, organ, etc. (emerge due to incremental buildup).
- **Robustness:** Allows system to survive and respond to change in inputs by buffering. Systems produced by evolution need to be robust to survive.

There are many methods of achieving robustness in such systems, including negative feedback, functional redundancy, multiscale redundancy, sparse coding (or compression to a few meaningful states), and Shannon's redundancy. All but Shannon's redundancy require a significant amount of energy. Multiscale redundancy is useful in that systems have evolved mechanisms to stop cascades of errors to multiple scales. Thus, for a mistake to propagate up to the emergent level, it must be made at multiple levels. In addition, all systems achieve robustness by hiding information. Examination of output states does not generally reveal information about internal states, which means systems are impossible to control by controlling the output. However, it also would not work to control a system at the level of fundamental mechanisms due to the number of error correction systems between fundamental mechanisms and outputs.

The idea of robustness of emergent systems can be used as a way to speculate about how these ideas may be applied to cancer. Viewing a patient and cancer as an emergent system uses Ashby's Law of Requisite Variety, the theory that a successful control system has to be as complex and have as many degrees of freedom as the system it is controlling. Although the consequences of information hiding make robust emergent systems hard to understand in detail, they can be easy to control and can be manipulated without much detailed understanding using the intermediate levels.



In approaching a cancer patient, a Markov model can be applied; this model system can be used for a robust, nonrandom system and has many abstract pseudostates (not real substates). The system can be modeled using probability distributions of the pseudostates. As an example, a cancer patient can be viewed as comprising three **control systems:**

- The patient's body
- The cancer (mutated from body control system, with different control systems)
- The patient/treatment loop

The idea is to help emergent control system 1, the body, win the battle over emergent control system 2, the cancer. Due to evolution, the body probably has more robustness (functional redundancy) than the cancer system; this robustness can be used as leverage. Control system 3, the treatment, can be used to help the body wrestle control from the cancer.

Keeping in mind that the purpose is to direct a treatment, the system can be viewed as a communication system with two channels—the communication channel from patient to physician

and the channel from physician to patient. The idea is to optimize the bandwidth in each channel. How the information is measured, what measurements are made in a patient, how the measurements are interpreted given the information, and what message is sent back to the patient given the message are key. Recall that information content is dependent on who interprets the information for what purpose. Information is not an absolute measure but is dependent on perspective. As pointed out by Shannon, when trying to maximally encode information in a channel, humans often put in levels and redundancy of meaning that can get in the way and be constraining. Encoding works most efficiently by ignoring most levels of meaning, avoiding redundancy, and using optimized code words to induce desired communication states, thereby optimizing the information bandwidth of the patient/treatment loop.

In the context of our model, this means determining whether there are extraneous levels of meaning. For example, skip the "disease" level; the "kind of cancer" is irrelevant in the treatment setting between the doctor and patient (although it may be useful for physician-to-physician communication). Instead of diagnosing the disease, treat the disease. In moving away from the traditional paradigm, information should be measured in orthogonal predictors, applying an ensemble of correctors. Given the combination of states, determine what combination of effectors will push the patient state to a point where the patient is more likely to wrestle control from the cancer. For example, instead of looking for biomarkers of disease in a proteomics scan, look at the ensemble of messages as an information code. What are the messages telling us about what treatment choices we should use?

Information theory is a useful tool for revealing what calculations to perform (i.e., determining the most informative things in an ensemble of measurements). Shannon's theory suggests that it is very likely that high-information indicators do not correspond nicely to understandable concepts; usually the code words are not what one is talking about. An analogous story may hold true for treatments. The high-information effectors probably do not correspond directly to understandable patterns. Highly targeted treatments may be the wrong tool. Treatments with more degrees of freedom, ensembles of treatments, or cocktails with many effects may be better, although difficult to put into practice.

In summary, fundamental understanding is worth pursuing and always helps, but understanding does not always translate into the levels needed for effective treatment. While different levels for interventions are often needed, keep in mind that what is understood is often not the best level to control. Examples of this approach indicate that use of specific biomarkers and treatments may not be the best therapeutic strategies.

Discussion Highlights: In the question-and-answer period, some of the concepts introduced by Dr. Hillis were further clarified. First, while Dr. Hillis stated that he does not know enough about cancer (i.e., what states correspond to in the system) to specify the most appropriate level for use of the Markov model, he does think that the Markov model could be applied. Dr. Hillis also clarified that he is not arguing against the use of biomarkers, but against specific diagnostic biomarkers. He is suggesting a new approach, a slightly different definition of the appropriate use of biomarkers. The current definition of informative biomarkers corresponds to a specific treatable disease state, and while sometimes correspondence may be demonstrated, this is too narrow a definition. If the definition of informative is changed to be in conjunction with all other biomarkers, can it give you the information to help create the right treatment cocktail? He suggests the information is there, but not in the form we are looking at. It was noted by others that real progress and payback have been obtained with some specific biomarkers (HER2/*neu*, ER) in informing us about diseases and treatments without detailed understanding of why the biomarkers are revealing. Furthermore, the absence of biomarkers has limited the usefulness of big prevention studies.

From the perspective of looking at evolution for new treatments from natural products and a recommendation to NCI to reevaluate natural products, Dr. Hillis was asked how he would test evolution for ways of controlling cancer. He suggested one approach of searching for emergent phenomena, or common points to many evolutionary systems, that would act as control points that could be exploited. Are there universal rules in evolutionary systems that can be used? Typically, we have looked at evolution from the standpoint of analysis rather than control. In response to a question

about adaptive systems, he noted that attacks on multiple fronts are often more successful than sequential hits.

A two-part question asked whether, given the work and information obtained so far, we should step back and look at where we are with cancer as a system, and what level should we be looking at now, given the information we have? Dr. Hillis suggested two approaches. The first is to put resources into obtaining more carefully controlled measures of disease (e.g., proteomics, 2D gels, biomarkers) and determine whether there are signal points. Second, it would also be worthwhile to do some pathfinding (i.e., to get high-leverage ideas of places where one would direct the "infantry"). A further need is to bring together disparate scientific groups to get better common understanding of systems, cell biology, and the perturbations employed in treatments.

Brainstorming Session: Elements for Addressing the Big Questions on Information and Communication in Cancer

Mr. Mittman led the group in this final brainstorming session to consider the elements that could address big questions in relation to information and communication in cancer. First, the big questions were defined, followed by suggestions of approaches to addressing the questions.

Group Discussion
Information in Cancer

What are the contextual and theoretical definitions of information in cancer? What are the components of information in cancer? How do we define information at the various levels of scale? The group posed the following approaches to answering these questions:

- Contextual information matters; the big question is how to define these broader states.
- Definition is context dependent, and there are complementary definitions depending on the process and scale (cells, tissues, patients, etc).
- What is the pertinent information? Due to the enormous number of variables, methods are needed to search feature space and find applicable variables. For the information relevant to question X, the answer can usually be obtained with a small number of variables.
- We need to define what space we are working on. Genome and protein spaces are probably wrong, but think about combinations of parameters and what they reflect for combination treatments. For example, in imaging, Fourier space is more productive than object space, as every object contributes to parameters. What is the best way to define the Achilles' heel in the system?
- Mathematically, we must define a probability distribution reflective of some biological function.
- Simplify as a measure of the number of choices (at a number of states).
- Consider that information changes at specific splice junctions.
- We need a "humanome," beyond genocentric, to capture global measurements of normal states and responses.
- Biological information, for example, surrogate markers, has resulted in good correlations and applications for disease treatment; biomarkers at individual protein levels do have value.
- In contrast to feature selection (which is artificially imposed), more appropriate questions would be: What are all variables/statistics (how many states can a cell be in?)? What does it take to specify the features? What feedback is required to understand variables? What meaningful states can the cell take on?

- Information may be in metagenes (composite measures of a large number of markers). Think about how we can measure markers of metagenes.
- We need knowledge, not necessarily information. Knowledge is the parts of the patient and the optimal conditions.

Group Discussion

Communication in Cancer

How is information communicated in cancer? What are the channels for information communication? The following observations were made:

- The only information used currently for cancer diagnosis is the pathologist's reading of the tissue sample. The information is in the cell shape; decoding and communication of diagnosis is with the pathologist. The emergent level is the structure.
- We need common data elements for a common language for information flow among physician specialties. (For example, a rules committee for pathologist reading is in process.)
- We need to start at the tissue level and move up to the point where mechanical properties are emergent.
- Cancer cells just want to proliferate; this is a fundamental principle of cancer; what are barriers to growth?
- We should look at the levels where we have won, for example, the protein (e.g., kinases), cell, and molecule (e.g., BCR-ABL) levels, not the tissue level.
- It is also important to look at the broader picture in the microenvironment and at the tissue level. This is where roadblocks are put in the way by limiting critical information to molecular biomarkers and genes.
- Look also at normal cells and physical forces.
- We need ways to measure communication happening through structural pathways.
- We need quantifiable physical information of cell/environmental interactions, network architecture, gradients, force, etc.
- As cancer is progressive, temporal measurements are also important, not just snapshots.

Breakout Session

A "Tour" of the Coding, Decoding, Transfer, and Translation of Information in Cancer: Defining the Scope of the Big Questions (Grand Challenges) and How To Approach Answering Them Through Transdisciplinary Research

The participants separated into four subgroups for the last breakout session and moved from station to station to discuss information in cancer research from four perspectives: (1) identifying critical information; (2) communication in cancer at multiple scales; (3) technology, models and tools; and (4) major overarching questions. The subgroups were tasked with providing input to NCI from these perspectives to assist in research planning. The breakout groups were asked to prioritize research questions among those already posed and select two that were of highest priority, list research strategies to answer the questions, and give the expected payoffs for cancer research from answering these questions. The discussion and reports from each group are summarized below.

Breakout 1: Information in Cancer

Chair: Wallace F. Marshall, Ph.D., Assistant Professor, University of California, San Francisco

Assuming the importance of information in cancer, group discussion focused on specific strategies to optimize using information to manage cancer, rather than defining information in cancer. Which theory was most appropriate to apply was also briefly discussed; this was recommended for more in-depth discussion at a future venue.

Discussion Highlights		
What Is Information in Cancer?		
Top Two Research Questions	***Top Research Strategies***	***Expected Payoffs***
1. What is the information that exists between the environment and cells? **2. What is the minimal sufficient model for cancer cells using information?**	**Quantify heterogeneity and dynamics.** **Iteratively incorporate new information and evolve models.** **Incorporate new types of information into models (e.g., mechanical).** **Requires interdisciplinary teams and new tools to test the predictions.**	Enables development of rational links between biomarkers and the outcome of treatment and course of the disease over time. Contributes to an evolving mechanistic understanding of the disease.

For a full graphical representation of this session, see Appendix 1.

A key question is what information exists between the environment and cells. To investigate this question, new information could be incorporated into models and new comprehensive models could be developed for prediction and testing (e.g., incorporating biomarkers). This would in turn require interdisciplinary teams and new tools to test the predictions and quantify heterogeneity and dynamics. The payoff would be development of rational links among biomarkers, treatment outcomes, and disease progression over time. This approach could be applied iteratively to develop better mechanistic understanding of the disease and hence better models. A key result would be to define the minimally sufficient model using information. Another major factor in the model is incorporating the environment, in which information constantly changes.

Breakout 2: Communication in Cancer at Multiple Scales

Chair: Brian Reid, M.D., Ph.D., Full Member, Divisions of Human Biology and Public Health Sciences, Fred Hutchinson Cancer Research Center

Strategies recommended to investigate communication in cancer across scales focused on use and development of appropriate new technologies in in vivo and model systems most closely mimicking the organism's complexity.

One key question is how to measure communication parameters as close to the in vivo state as possible using new technologies to measure cell parameters, the microenvironment, and metabolism (e.g., the Warburg effect). Research strategies would include making measurements at multiple levels and integrating the results. Although single-cell measurements are very important, cell population dynamics

and microenvironment effects are critical to understanding the ecology of the cancer system. It would be useful to examine past successes and failures using these concepts to describe cancers.

Discussion Highlights		
Cancer Communication Across Scales		
Top Two Research Questions	***Top Research Strategies***	***Expected Payoffs***
1. How do we measure communication parameters as close to the in vivo state as possible (using new technologies to measure cell parameters, microenvironment, and metabolism (e.g., the Warburg effect)? **2. How does cancer kill its human host (recurrence)?**	**Measure ecology of the cancer system at multiple levels (cell, microenvironment, metabolism) incorporating new technologies as needed; integrate the results.** **Develop/apply emerging in vivo measurement technologies and 3D in vitro cell culture systems.** **Study precancer, including nuclear morphology, by imaging, signaling pathways, accumulation of genetic aberrations, system adjustments, tissue morphometry.** **Develop collaborations to bring physical scientists into picture.**	Control cancer (cure). ▪ Composite measure (metagene) ▪ Composite medication (cocktail) Better diagnostics, localization for early detection. Faster pace of research. Predict outcomes. Learn to control cancer.

For a full graphical representation of this session, see Appendix 1.

A second key question considered was how does cancer kill its host. To address this question, research could start with hollow-organ cancers, using nanotechnology methods and biopsies to investigate the course of cancer in the cell, microenvironment, and metastases. In addition, intermediate-scale models could be developed; composite measurements could be made in these models, and composite interventions could be used to elicit outcomes for interventions. For example, tumors could be classified according to response of models to drug treatment, and premalignant lesions could be classified by measuring risk for progression. These models would be developed iteratively and adaptively. The payoffs for research addressing these questions would be new cancer control strategies involving complex measurements for cancer risk and composite intervention (drug cocktails). This work would result in better diagnosis, a faster pace of research from iterative adaptation of strategies, and improvement in predicting outcomes, particularly in precancerous states.

Breakout 3: Technology, Models, and Tools

Chair: Thomas V. O'Halloran, Ph.D., M.A., Professor, Northwestern University

Group discussion addressed developing appropriate models and tools ranging from experimental to mathematical approaches and use of databases.

Discussion Highlights

Technology, Models, and Tools

Top Three Research Questions	*Top Research Strategies*	*Expected Payoffs*
1. Compare cancer and normal states; what are the properties that distinguish the cancer? **2. How much energy is expended in cancer evolution?** **3. What are the rules that govern cancer evolution?**	**Develop new methods to precisely measure phenomena such as elasticity, chemical gradients, in vivo dynamics, nucleosome localization (using microscopy), multiple signals/ responses simultaneously.** **Deconvolution of heterogeneity →single-cell resolution→back to emergent property of tumor.** **Identify time-dependent order parameters from database analysis.** **Develop multiscale dynamic models that include important global variables to provide predictive/time-dependent computational physics-based simulation schemes.**	Ability to relate measurements to outcome of therapy. With the genotype focus would have the ability to: ▪ Identify new order parameters that could permute to the clinic. ▪ Relate models/fluxes to current drugs (e.g., those measured in NCI 59 cell line database).

For a full graphical representation of this session, see Appendix 1.

Using new models and tools could help address key questions on the properties that distinguish cancer from normal states, the amount of energy expended in cancer evolution, and the rules that govern cancer evolution. Each of these questions can be approached by collecting data with multiple types of measurements, using those data to develop phenotypic prognostic parameters, then developing test models at all scales to relate the parameters to treatment outcomes. Particularly interesting is development of new methods to precisely measure parameters such as elasticity, chemical gradients, multiple dynamic signaling, nucleosome location (including integrated views, measurements of systems, measurements in different cells, etc.), and cellular energy expenditure (e.g., define energy budgets for systems before and after metastases). To examine metabolism and energy, energy should be measured before and after invasion. Imaging tools such as MRI or fluorescence temperature-sensitive probes might assist this effort. A central repository, perhaps at NCI, for all types of data used to develop models would be useful. It was noted that patient privacy policies (Health Insurance Portability and Accountability Act [HIPAA]) complicate making human data available through a repository. This issue could be addressed by developing an authentication scheme for investigators (e.g., with institutions taking responsibility for protecting the data for their investigators, such as outlined in the U.S. Department of Energy [DOE] Human Subjects Protection Program). This would lead to faster discovery of emergent tumor properties and better understanding of relationships between parameters at all scales that affect cancer, leading to better treatments for disease. An example of these relationships is the link between p53, mitochondrial shape, and cancer. Developing models that associate levels of cell and organism would accelerate the pace and reduce the cost of research.

Breakout 4: Major Overarching Questions

Chair: Carlo C. Maley, Ph.D., Assistant Professor, The Wistar Institute

The group discussion focused primarily on defining patient metastates and cellular architectural changes to improve control of cancer.

Discussion Highlights

Major Overarching Questions

Top Two Research Questions	*Top Research Strategies*	*Expected Payoffs*
1. What are the metastates of the cancer and patient? **2. What causes cell/architecture changes?**	**Define quantitative measures of metastates.** **Collect quantitative data on patient-derived samples:** ▪ Pathological variables ▪ Imaging ▪ Clinical ▪ Genetic ▪ Proteomics, etc. [-omics] **Use modeling; include normal controls; collapse state space→look for clustering; develop reproducible measures.** **Define causes of cell/architecture changes using 3D cultures; time course; exploration of initial conditions; development of dynamic cellular proteomics in vitro and in vivo; computations/math modeling.**	New treatment modes→how to change metastates. Integrated understanding of cancer. New tools for prognosis and diagnostic and therapy management. Connect molecular biology and pathology. Therapies that normalize the tissue. Response map for cells and tissues. Develop a control theory for managing cancer. Understanding cancer and tissue dynamics. Understanding interactions of patient systems (e.g., immunology).

For a full graphical representation of this session, see Appendix 1.

The first overarching question is what metastates describe the behavior of the system (cancer and patient)? To identify metastates, all possible data on normal and disease states would be collected from multiple sources; then data would be collapsed into metavariables, looking for clustering in order to define quantitative measures. This approach requires good-quality (low noise), reproducible measurements. Patient-derived data, such as pathological variables, imaging, and clinical, genetic, and proteomic measurements, would be collected, along with data from experimental systems paired with modeling. This approach could yield new cancer control strategies. These control strategies would be intended to maintain patients in stable metastates for the long term (but would not necessarily kill all the cancer cells); they would allow practical management of cancer. The approaches would result in new treatment modes to change the metastates and a control theory for managing cancer, as well as new tools for diagnosis, prognosis, and therapy management.

A second question considered was what causes cellular and architectural changes in cancer and neoplastic progression? To investigate, both patient and experimental systems would be used, focusing on time courses of state changes. Further clarification is needed in control of chromosome amplification and rearrangement in evolution and cancer. Also, examining evolution of infectious disease may be informative, and game theory would be an exciting tool to use. Other techniques

could include dynamic cellular in vitro and in vivo proteomics and 3D cultures, complemented with modeling. Investigations of architectural changes would permit further understanding of connections between molecular biology and pathology, cancer, and tissue dynamics. Determining the possibility of manipulating stem cell differentiation by controlling the microenvironment could be very useful. All these strategies could allow evaluation of interventions that normalize tissue.

For a full graphical representation of this session, see Appendix 1.

Summary and Next Steps

Anna D. Barker, Ph.D., Deputy Director, NCI, and John E. Niederhuber, M.D., Director, NCI

Dr. Barker thanked the attendees for their participation and contributions, particularly those who gave keynote and shorter panel presentations, and Mr. Mittman for his excellent meeting facilitation. She thanked Dr. Niederhuber for his support of this innovative and potentially paradigm-changing initiative, pointing out that this was not always easy for the NCI Director in tough budget times. She noted that, like the two prior think tanks, this meeting had exceeded expectations. This group has set a high bar for exploration of the very complex areas of information and its management in cancer. The questions posed in all the sessions were important to developing a fundamental understanding of cancer, and many offer new approaches to ultimately controlling the disease. Interestingly, this meeting posited that we may need to look more closely at the level and depth of information required for control of the disease.

Dr. Niederhuber also voiced his thanks for the participants' contributions. The dialogue and format of the meeting were stimulating and enriching. The fact that so many participants told him that, as a result of these meetings, they think about their work differently and have established new relationships with colleagues makes the meeting even more worthwhile. It is his goal to keep the momentum from this and the prior think tanks moving ahead. Dr. Barker will present a proposal to the NCI Board of Scientific Advisors to move forward with a funding instrument to allow support for a network of centers to pursue these innovative new directions. Dr. Niederhuber and Dr. Barker plan to continue this series of think tanks in the following year, as the science will move rapidly and there are a number of areas yet to explore. Dr. Niederhuber adjourned the meeting.

Appendix 1. Meeting Sketches

Figure 1. Keynote Presentation
Is DNA a Molecule? Musings on Good Cells Making Bad Choices
Robert Phillips

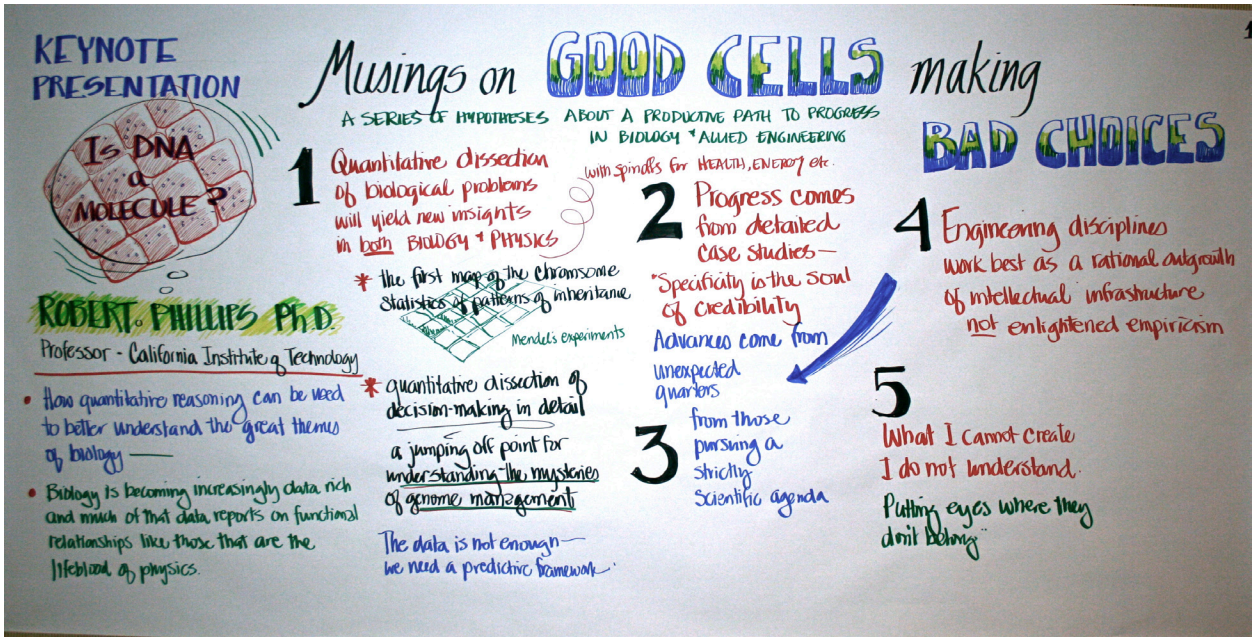


Figure 2. Welcome and Introduction of Keynote Presentation
John E. Niederhuber

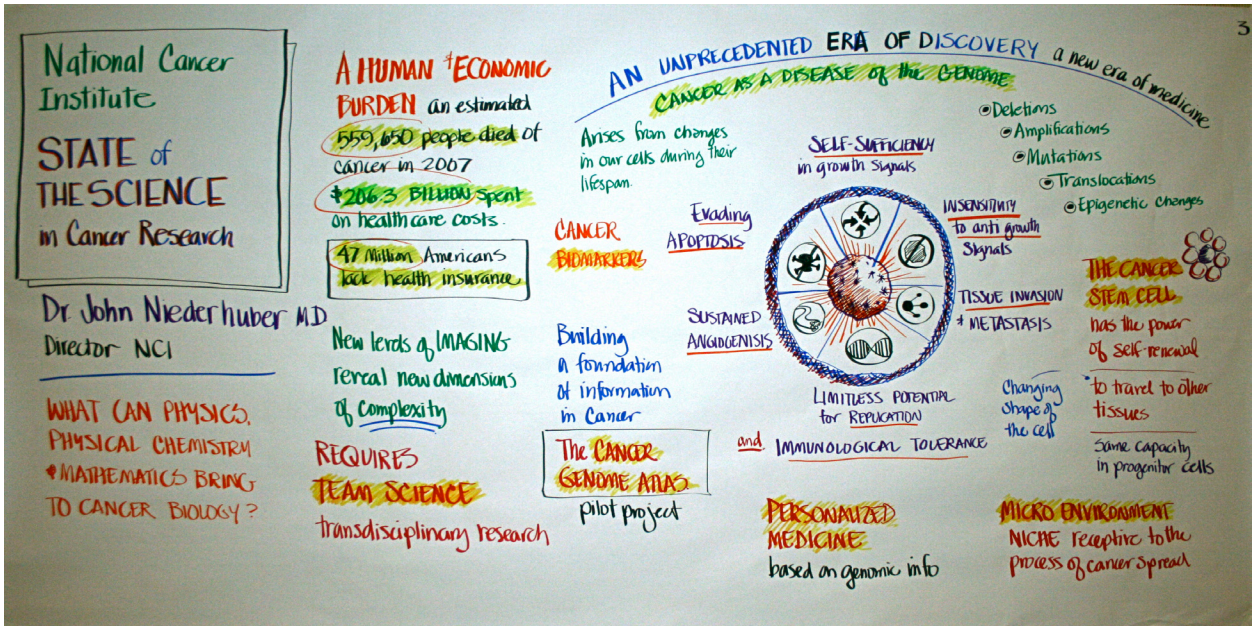


Figure 3. Keynote Presentation
Information Theory in Molecular Biology: Key to Understanding Information Transfer, Signaling, and Translation in Cancer
Christoph C. Adami

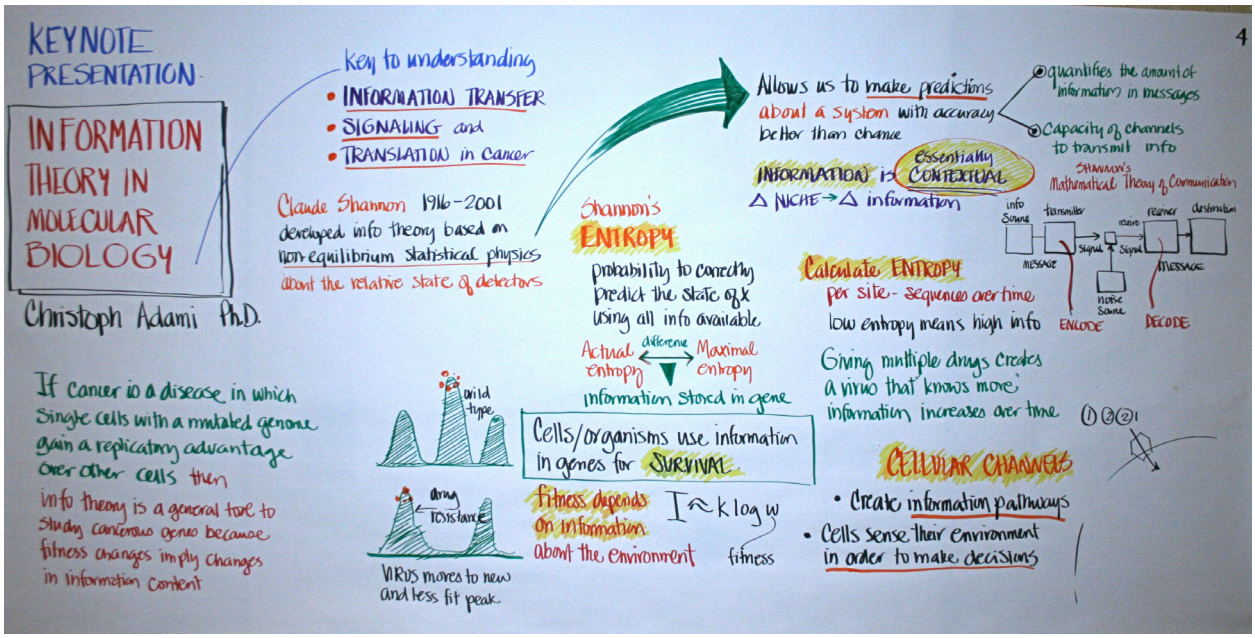


Figure 4. Keynote Presentation
The Information: Genetic Code(s) and Cancer—State of the Science
David Haussler

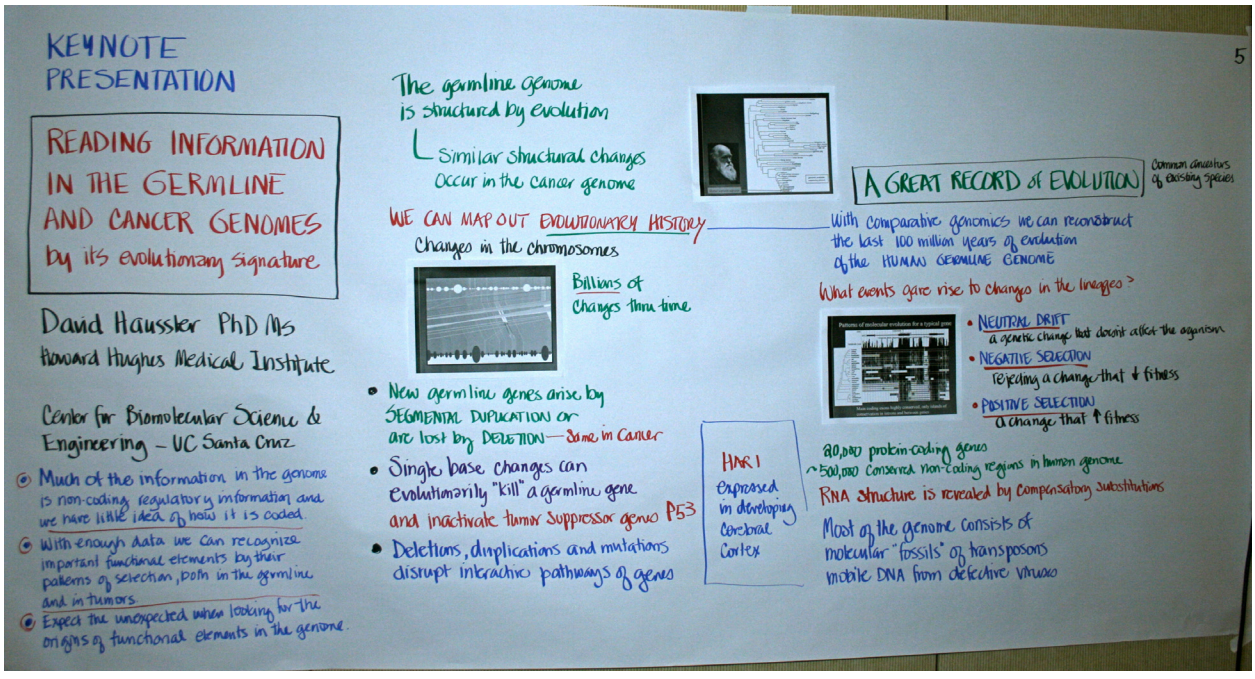


Figure 5. Keynote Presentation
The Rest of the Story: The Small RNAs and Cancer
Phillip A. Sharp

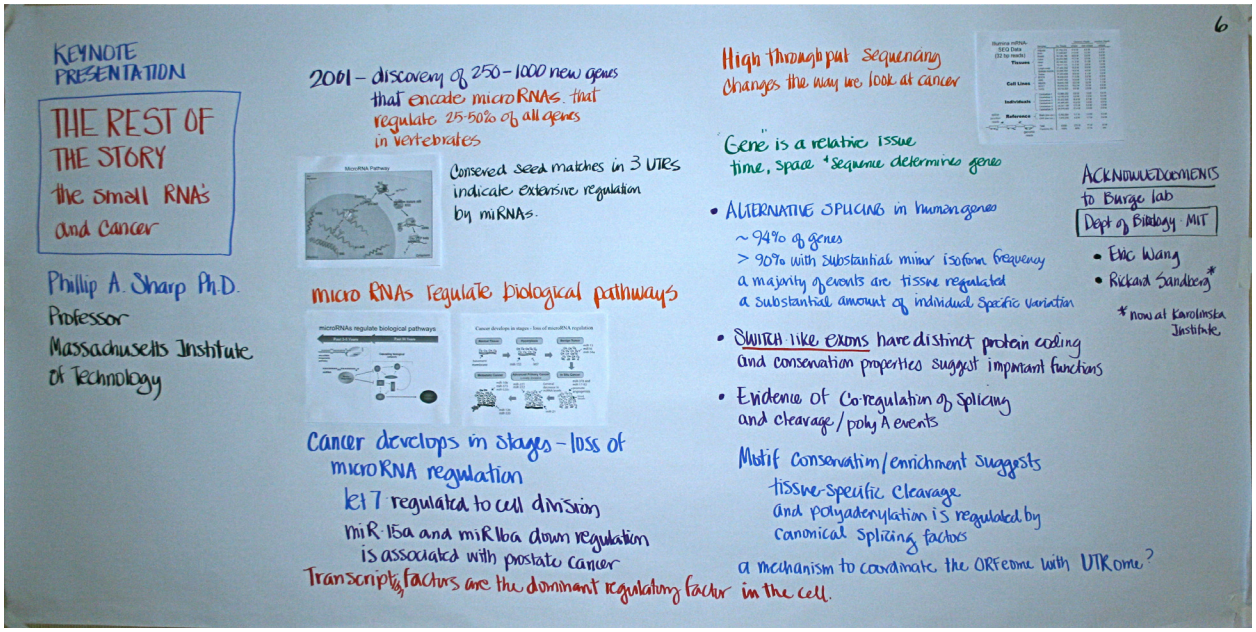


Figure 6. Small Group Discussion
Information Theory—If It's So Important in Cancer, Why Have We Not Made More Progress in the Field?
Robert Mittman and Group

Figure 7. Brief Presentations
Contextual Translation of Information: So Many Signals, So Many Channels, So Much Translation on So Many Scales
Darryl K. Shibata, Philip R. LeDuc, Mauro Ferrari, Jennifer Lippincott-Schwartz, Robert A. Gatenby

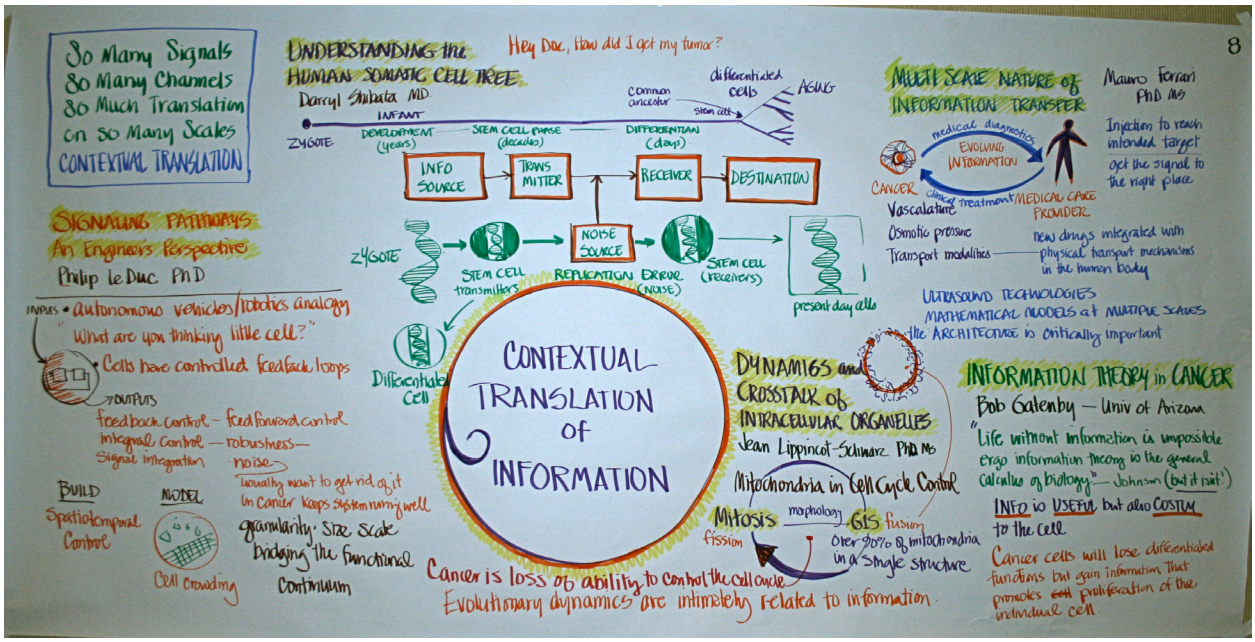


Figure 8. Small Group Discussions
Understanding Signaling and Contextual Translation of Information at Multiscales: What's Relevant From the Physical Sciences?
Facilitator: Robert Mittman

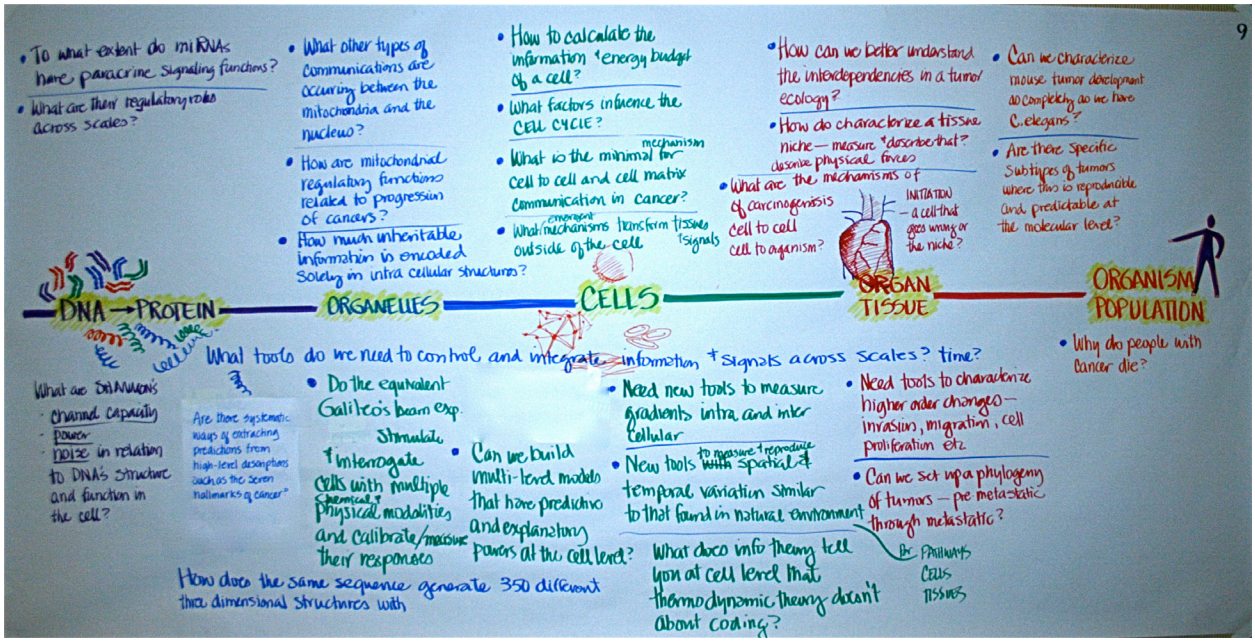


Figure 9. Panel Discussion
The Outcomes and Consequences of Information Transfer in Cancer Across Length Scales
Wallace F. Marshall, Carlo C. Maley, Robert H. Austin, Christoph C. Adami

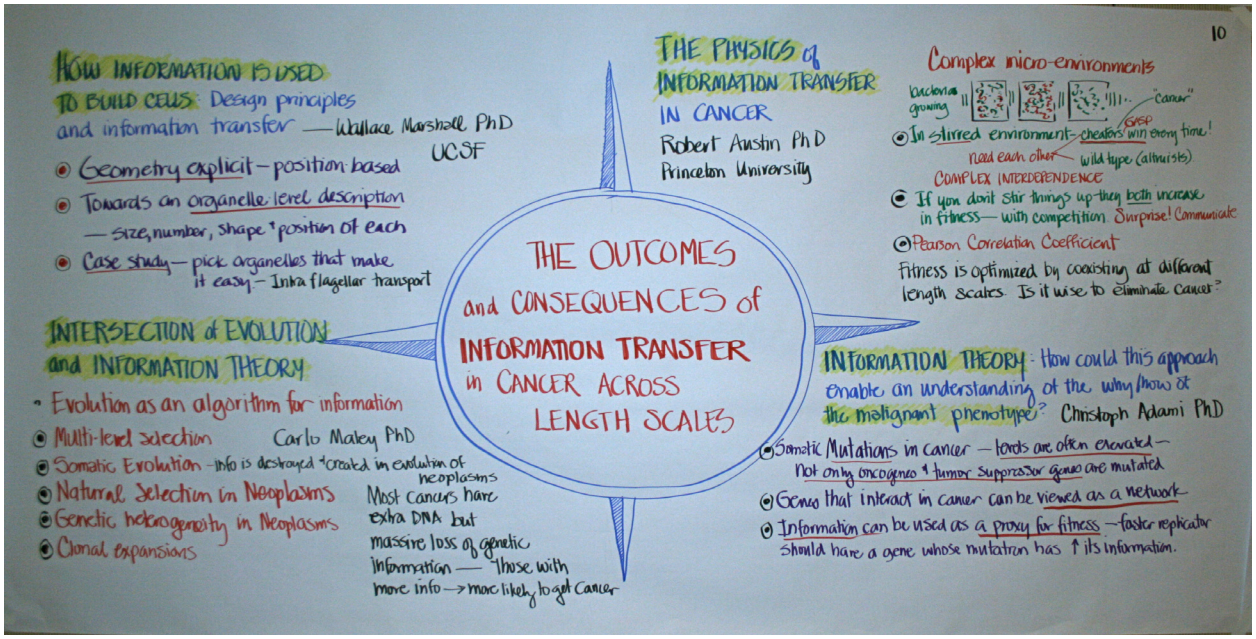


Figure 10. Panel Discussion
The Future: If We Understand the Specifics (Physics, Chemistry, etc.) of the Information, Its Transfer, and Contextual Translation at Multiple Length Scales in Cancer, Can We Alter Outcomes?
Paul Davies, Donald S. Coffey, Robert Phillips, W. Daniel Hillis, John E. Niederhuber

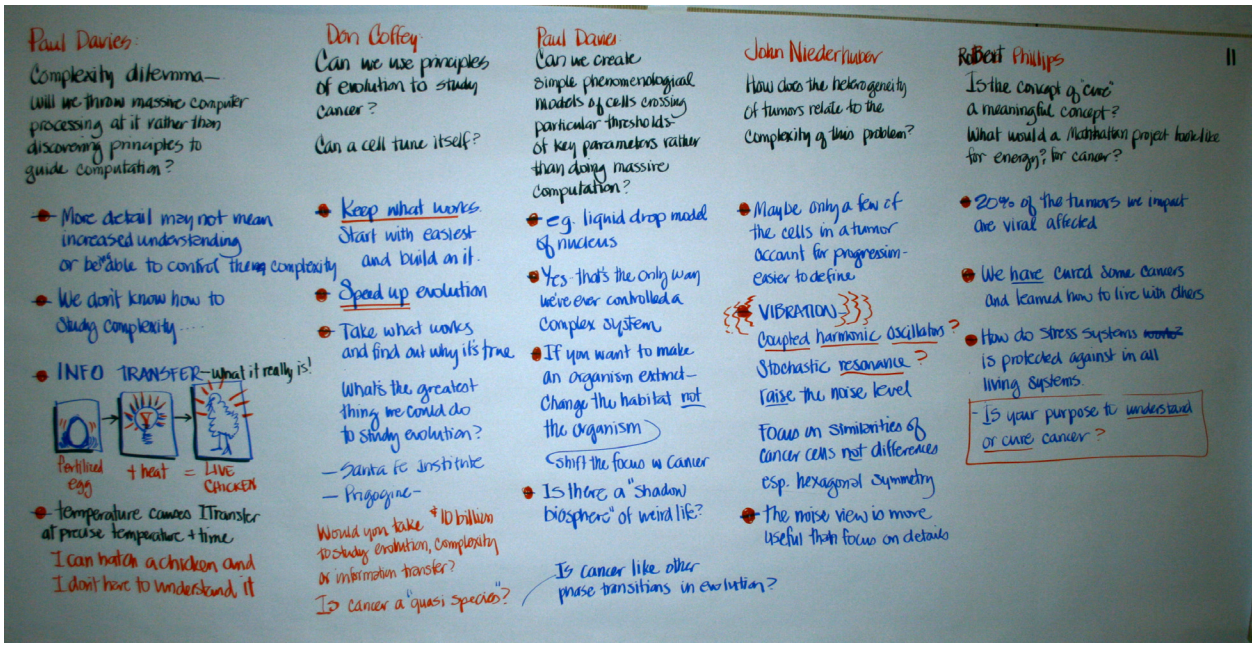
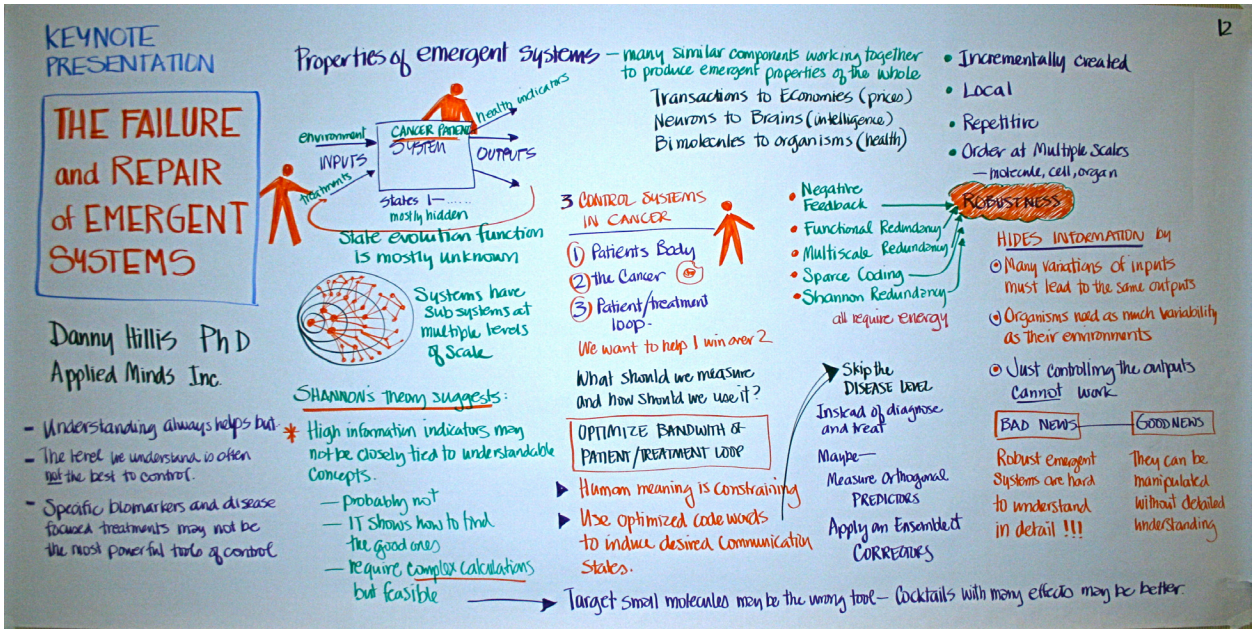


Figure 11. Keynote Presentation

The Failure and Repair of Emergent Systems: A Systems Engineering Approach to Cancer

W. Daniel Hillis



Appendix 2. Bibliography

1. Gerlich, D., J. Beaudouin, et al. (2003). "Global chromosome positions are transmitted through mitosis in mammalian cells." Cell 112(6): 751-64.

2. Kirchhamer, C. V., L. D. Bogarad, et al. (1996). "Developmental expression of synthetic cis-regulatory systems composed of spatial control elements from two different genes." Proc Natl Acad Sci U S A 93(24): 13849-54.

3. Poirier, M. G., M. Bussiek, et al. (2008). "Spontaneous access to DNA target sites in folded chromatin fibers." J Mol Biol 379(4): 772-86.

4. Mumm, J. P., A. Landy, et al. (2006). "Viewing single lambda site-specific recombination events from start to finish." Embo J 25(19): 4586-95.

5. Small, S., A. Blair, et al. (1992). "Regulation of even-skipped stripe 2 in the Drosophila embryo." Embo J 11(11): 4047-57.

6. Alberts, B. (1998). "The cell as a collection of protein machines: preparing the next generation of molecular biologists." Cell 92(3): 291-4.

7. Bissell, M. J., and M. A. Labarge (2005). "Context, tissue plasticity, and cancer: are tumor stem cells also regulated by the microenvironment?" Cancer Cell 7(1): 17-23.

8. Chin, S. F., A. E. Teschendorff, et al. (2007). "High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer." Genome Biol 8(10): R215.

9. Cancer Genome Atlas Research Network. (2008) ."Comprehensive genomic characterization defines human glioblastoma genes and core pathways." Nature 455(7216): 1061-8.

10. Morin, R. D., M. D. O'Connor, et al. (2008). "Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells." Genome Res 18(4): 610-21.

11. Lee, R. C., R. L. Feinbaum, et al. (1993). "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14." Cell 75(5): 843-54.

12. Wightman, B., I. Ha, et al. (1993). "Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans." Cell 75(5): 855-62.

13. Lagos-Quintana, M., R. Rauhut, et al. (2001). "Identification of novel genes coding for small expressed RNAs." Science 294(5543): 853-8.

14. Lau, N. C., L. P. Lim, et al. (2001). "An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans." Science 294(5543): 858-62.

15. Lee, R. C., and V. Ambros (2001). "An extensive class of small RNAs in Caenorhabditis elegans." Science 294(5543): 862-4.

16. Bonci, D., V. Coppola, et al. (2008). "The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities." Nat Med 14(11): 1271-7.

17. Jacob, F., and J. Monod (1961). "Genetic regulatory mechanisms in the synthesis of proteins." J Mol Biol 3: 318-56.

18. Shannon, C. E. (1948). "A mathematical theory of communication." Bell System Technical J 27: 379-423, 623-656.

Appendix 3. Meeting Agenda

Overview

This is the third in a series of NCI "think tanks" that bring together leaders from the physical sciences with basic and clinical cancer researchers to explore approaches that may contribute to solving intractable problems that we face in understanding and controlling cancer. Although the conversations in the first meeting identified a large number of potential research opportunities, four major themes emerged for further exploration as follows: the "physics" of cancer (e.g., forces and mechanics, thermodynamics, gradients, etc.); evolution and evolutionary theory in cancer; information coding, transfer translation, and information theory in cancer; and the complexity of cancer.

The second meeting in this series focused on "A New Look at Evolution and Evolutionary Theory in Cancer." This meeting identified a number of the major research questions in the field and elaborated a number of "grand challenges" that, if met, would significantly improve our understanding of the role of evolution in cancer. Underlying many of the conversations at this think tank were questions on the role of information and information theory in cancer, specifically those changes that confer selective advantages. Overall it was clear that a great deal of knowledge is needed to elucidate the role of information flow at all scales in understanding the emergence of the malignant phenotype.

Although this think tank will focus on the coding, decoding, flow, and translation of information in cancer, our conversations will by necessity reflect in an integrative way all four of the themes that derived from the first meeting. Our overall goal for this meeting is to better define and understand this complex field relative to its potential role in understanding and controlling cancer. Overall we plan to:

- Explore the concept of what "information" means in terms of the genetic code and its translation in cancer relative to context and certain specific aspects that characterize cancer.
- From the perspective of both the physical and biological sciences, determine the "state of the science" of information and information theory in terms of understanding cancer at all scales.
- Identify the major critical research questions in the state of the science of information and information sciences in cancer that could represent major areas for transdisciplinary research.
- Determine where/how innovative research approaches in information/information theory might lead to the development of new cancer interventions.
- Offer guidance on how the NCI can integrate areas from the physical sciences (physics, mathematics, chemistry, engineering, etc.) with cancer biology/oncology to enable the development of this field of study.

Outcomes

It is anticipated that the outcomes of this think tank will enable the development of the innovative strategies, models, and approaches needed to build this transdisciplinary field of cancer information coding, transfer, and translation as well as its theoretical foundation. Input from the meeting will be utilized to inform new research directions and mechanisms that will hopefully energize and advance this convergent field of cancer research. Specifically, targeted outcomes include the following:

- Produce a detailed view and interpretation of the state of the field of information and information theory related to cancer.
- Assuming that the field is not currently a major thrust in terms of our research efforts to understand and control cancer, define the barriers that are limiting the development of the field.

- If progress in the field of information theory and information management applied to cancer is to be achieved in a timely manner, define major research questions and directions for the future.
- Propose examples of research strategies, data management approaches, and infrastructure that could be employed to inform and support addressing these research questions.

The conversations comprising this think tank, including brainstorming sessions, presentations, roundtables, and reports from work groups, will be captured in a report that will be available on an NCI Web site dedicated to this Physical Sciences-Based Frontiers in Oncology Series.

Agenda

Wednesday, October 29

5:00 p.m. - 6:00 p.m.	**Registration**	*Salon III Foyer*
6:00 p.m. - 7:15 p.m.	**Reception and Buffet Dinner**	*Salon III*
7:30 p.m. - 7:50 p.m.	**Meeting Background and Introductions** Anna D. Barker, Ph.D. Deputy Director National Cancer Institute **Welcome and Introduction of Keynote Presenter** John E. Niederhuber, M.D. Director National Cancer Institute	
7:50 p.m. - 8:50 p.m.	Keynote Presentation ***Is DNA a Molecule? Musings on Good Cells Making Bad Choices*** Robert Phillips, Ph.D. Professor California Institute of Technology ***Questions/Discussion***	
8:50 p.m. - 9:00 p.m.	**Think Tank Process** Anna D. Barker, Ph.D. Deputy Director National Cancer Institute	
9:00 p.m. - 9:10 p.m.	**Process and Outcomes Overview** Facilitator: Robert Mittman, M.S., M.P.P. Founder/President Facilitation, Foresight, Strategy	

Thursday, October 30

7:00 a.m. - 8:00 a.m. **Continental Breakfast**

8:00 a.m. - 8:30 a.m. **The NCI's Physical Sciences-Based Frontiers in Oncology Think Tank Series** *Salon III*
Anna D. Barker, Ph.D.
Deputy Director
National Cancer Institute

Think Tank Process
Facilitator: Robert Mittman, M.S., M.P.P.
Founder/President
Facilitation, Foresight, Strategy

Welcome and Introduction of Keynote Presentation
John E. Niederhuber, M.D.
Director
National Cancer Institute

8:30 a.m. - 9:00 a.m. **Keynote Presentation**
Information Theory in Molecular Biology: Key to Understanding Information Transfer, Signaling, and Translation in Cancer
Christoph C. Adami, Ph.D.
Professor
California Institute of Technology

9:00 a.m. - 10:30 a.m. **Small Group Discussions: Information Theory – If It's So Important in Cancer, Why Have We Not Made More Progress in the Field?**
Facilitator: Robert Mittman, M.S., M.P.P.

10:30 a.m. - 10:45 a.m. **Break**

10:45 a.m. - 11:15 a.m. **Keynote Presentation**
The Information: Genetic Code(s) and Cancer—State of the Science
David Haussler, Ph.D., M.S.
Professor
University of California, Santa Cruz

11:15 a.m. - 11:45 a.m. **Keynote Presentation**
The Rest of the Story: The Small RNAs and Cancer
Phillip A. Sharp, Ph.D.
Professor
Massachusetts Institute of Technology

11:45 a.m. - 12:15 p.m. **Group Discussion: Cancer Information**
Dr. Adami, Dr. Haussler, Dr. Sharp, and Group

Facilitator: Robert Mittman, M.S., M.P.P.

12:15 p.m. - 1:10 p.m. **Lunch**

1:10 p.m. - 2:30 p.m. **Contextual Translation of Information: So Many Signals, So Many Channels, So Much Translation on So Many Scales**
Panel: Brief Presentations

Beyond the Genome: Understanding the Human Somatic Cell Tree
Darryl K. Shibata, M.D.
Professor
University of Southern California

Signaling Pathways: An Engineer's Perspective
Philip R. LeDuc, Ph.D.
Associate Professor
Carnegie Mellon University

Multiscale Nature of Information Transfer
Mauro Ferrari, Ph.D., M.S.
Professor
University of Texas Health Science Center at Houston

Dynamics and Cross-Talk of Intracellular Organelles
Jennifer Lippincott-Schwartz, Ph.D., M.S.
Senior Investigator
National Institute of Child Health and Human Development

Information Theory in Living Systems: Contributions of the Microenvironment
Robert A. Gatenby, M.D.
Division Chief
Moffitt Cancer Center and Research Institute

Discussion
Facilitator: Robert Mittman, M.S., M.P.P.

2:30 p.m. - 3:45 p.m. **Small Group Discussions: Understanding Signaling and Contextual Translation of Information at Multiscales: What's Relevant From the Physical Sciences?**

Facilitator: Robert Mittman, M.S., M.P.P.

****

3:45 p.m. - 6:00 p.m. **Mind-Clearing Break**

****

6:00 p.m. **Think Tank Reconvenes** *Salon III*

6:00 p.m. - 6:30 p.m. **Working Dinner**

6:30 p.m. - 7:30 p.m.

The Outcomes and Consequences of Information Transfer in Cancer Across Length Scales
Panel Discussion

How Information Is Used To Build Cells: Design Principles and Information Transfer (10-minute overview)
Wallace F. Marshall, Ph.D.
Assistant Professor
University of California, San Francisco

Intersection of Evolution and Information Theory: What Does It Mean for Cancer? (5-minute perspective)
Carlo C. Maley, Ph.D.
Assistant Professor
The Wistar Institute

The Physics of Information Transfer in Cancer (5-minute perspective)
Robert H. Austin, Ph.D.
Professor of Physics
Princeton University

Information Theory: Could This Approach Enable an Understanding of the Why/How of the Malignant Phenotype? (5-minute perspective)
Christoph C. Adami, Ph.D.
Professor
California Institute of Technology

Discussion
Facilitator: Robert Mittman, M.S., M.P.P.

7:30 p.m. - 8:30 p.m.

Small Group Discussions: From the Viewpoint of Information Transfer and Translation: New Research Approaches/Directions to Better Understand the Cancer Process at Multiscales
Facilitator: Robert Mittman, M.S., M.P.P.

8:30 p.m. - 9:30 p.m.

The Future: If We Understand the Specifics (Physics, Chemistry, etc.) of the Information, Its Transfer, and Contextual Translation at Multiple Length Scales in Cancer, Can We Alter Outcomes?
Panel Discussion

Paul Davies, Ph.D., D.Sc.
Professor
Arizona State University

Donald S. Coffey, Ph.D.
Professor
Johns Hopkins University

Robert Phillips, Ph.D.
Professor
California Institute of Technology

W. Daniel Hillis, Ph.D.
Chairman
Applied Minds, Inc.

John E. Niederhuber, M.D.
Director
National Cancer Institute

Discussion
Facilitator: Robert Mittman, M.S., M.P.P.

Friday, October 31

7:00 a.m. - 8:00 a.m.	**Continental Breakfast**	*Salon III*
8:00 a.m. - 8:15 a.m.	**Review of Day 1** Robert Mittman, M.S., M.P.P. Founder/President Facilitation, Foresight, Strategy	
8:15 a.m. - 9:00 a.m.	**Keynote Presentation** ***The Failure and Repair of Emergent Systems: A Systems Engineering Approach to Cancer*** W. Daniel Hillis, Ph.D. Chairman Applied Minds, Inc. ***Questions and Discussion***	
9:00 a.m. - 11:30 a.m.	**A "Tour" of the Coding, Decoding, Transfer, and Translation of Information in Cancer: Defining the Scope of the Big Questions (Grand Challenges) and How to Approach Answering Them Through Transdisciplinary Research** ***Thinking Groups***	*Salon II, Plaza B, Plaza D, and Diplomat*
	Individual Group Facilitation Facilitator: Group Leader Facilitators Robert Mittman, M.S., M.P.P.	
11:30 a.m. - 1:00 p.m.	**Reporting and Refining the Grand Challenges** **Group Reporting** Facilitator: Robert Mittman, M.S., M.P.P.	
1:00 p.m. - 1:30 p.m.	**Summary and Next Steps** John E. Niederhuber, M.D. Director National Cancer Institute Anna D. Barker, Ph.D. Deputy Director National Cancer Institute	

Appendix 4. Meeting Participants

Christoph C. Adami, Ph.D.
Professor
Keck Graduate School of Applied Life Sciences
535 Watson Drive
Claremont, CA 91711-4817
(909) 607-9853
(909) 607-8086 Fax
adami@kgi.edu

David B. Agus, M.D.
Director
Spielberg Family Center for Applied Proteomics
Cedars-Sinai Medical Center
Suite 215E
8631 West Third Street
Los Angeles, CA 90048
(310) 423-7620
(310) 423-1998 Fax
david.agus@cshs.org

Gaurav Arya, Ph.D.
Assistant Professor of Nanoengineering
Department of Mechanical and Aerospace Engineering
University of California, San Diego
261 EBU II
Mail Code 0411
9500 Gilman Drive
La Jolla, CA 92093-041
(858) 822-5542
(858) 534-5698 Fax
garya@ucsd.edu

Robert H. Austin, Ph.D.
Professor of Physics
Department of Physics
Princeton University
122 Jadwin Hall
Princeton, NJ 08544
(609) 258-4353
(609) 258-1115 Fax
austin@princeton.edu

Gang Bao, Ph.D.
Robert A. Milton Chair in Biomedical Engineering
Georgia Institute of Technology and Emory University
313 Ferst Drive, NW
Atlanta, GA 30332
(404) 385-0373
(404) 894-4243 Fax
gang.bao@bme.gatech.edu

Anna D. Barker, Ph.D.
Deputy Director
National Cancer Institute
National Institutes of Health
Building 31, Room 11A-30
MSC 2580
31 Center Drive
Bethesda, MD 20892-2580
(301) 496-1045
(301) 480-2889 Fax
barkera@mail.nih.gov

Ravi V. Bellamkonda, Ph.D.
Professor
Department of Biomedical Engineering
Georgia Institute of Technology/Emory University
313 Ferst Drive, NW
Atlanta, GA 30332
(404) 385-5038
(404) 385-5044 Fax
ravi@bme.gatech.edu

Sherrin Bennett, M.S.
President
Interactive Learning Systems
214 Water Street
Point Richmond, CA 94801
(510) 233-2230
sherrinbennett@earthlink.net

Carl T. Bergstrom, Ph.D.
Associate Professor
Department of Biology
University of Washington
Box 351800
Seattle, WA 98195-1800
(206) 685-3487
cbergst@u.washington.edu

Jordan D. Berlin, M.D.
Clinical Director
Gastrointestinal Oncology
Associate Professor of Medicine
Medical Oncologist
Vanderbilt-Ingram Cancer Center
777 Preston Building
Nashville, TN 37232-6307
(615) 322-4967
(615) 343-7602 Fax
jordan.berlin@vanderbilt.edu

Kenneth H. Buetow, Ph.D.
Director
Center for Bioinformatics and Information Technology
National Cancer Institute
National Institutes of Health
Suite 6000
2115 East Jefferson Street
Bethesda, MD 20892
(301) 435-1520
(301) 480-6641 Fax
buetowk@mail.nih.gov

Pedro Cano, M.D.
The University of Texas M.D. Anderson Cancer Center
Box 137
1515 Holcombe Boulevard
Houston, TX 77030-4095
(713) 792-6313
(713) 794-1824 Fax
pcano@mdanderson.org

Donald S. Coffey, Ph.D.
Professor of Urology Research
Johns Hopkins School of Medicine
Marburg Building, Room 121
600 North Wolfe Street
Baltimore, MD 21287-2101
(410) 955-2517
(410) 502-9336 Fax
dcoffey@jhmi.edu

Carolyn C. Compton, M.D., Ph.D.
Director
Office of Biorepositories and Biospecimen Research
National Cancer Institute
National Institutes of Health
Building 31, Room 10A-03
MSC 2580
31 Center Drive
Bethesda, MD 20892-2580
(301) 402-1762
(301) 480-4814 Fax
comptcar@mail.nih.gov

Peter T. Cummings, Ph.D.
John Robert Hall Professor of Chemical Engineering
Professor of Chemical and Biomolecular Engineering
Vanderbilt University
VU Station B 351604
303 Olin Hall
Nashville, TN 37235
(615) 322-8129
(615) 343-7951 Fax
peter.cummings@vanderbilt.edu

Paul Davies, Ph.D., D.Sc.
Professor and Director
The Beyond Center for Fundamental Concepts in Science
Arizona State University
P.O. Box 871504
Tempe, AZ 85287-1504
(480) 965-3240
(480) 965-7954 Fax
paul.davies@asu.edu

Michael W. Deem, Ph.D.
John W. Cox Professor of Bioengineering
Professor of Physics and Astronomy
Rice University
Mail Stop 142
6100 Main Street
Houston, TX 77005-1892
(713) 384-5811
(713) 348-5852 Fax
mwdeem@rice.edu

Micah X. Dembo, Ph.D.
Professor of Biomedical Engineering
Cellular and Subcellular Mechanics Laboratory
Boston University
44 Cummington Street
Boston, MA 02215-2407
(617) 353-1671
(617) 353-1671 Fax
mxd@bu.edu

Emmanuele DiBenedetto, Ph.D.
Centennial Professor of Mathematics
Professor of Molecular Physiology and Biophysics
Department of Mathematics
Vanderbilt University
Stevenson Center
Nashville, TN 37240
(615) 343-5906
(615) 343-0215 Fax
em.diben@vanderbilt.edu

Alexander G. Dimitrov, Ph.D.
Assistant Professor
Center for Computational Biology
Montana State University
P.O. Box 173505
Bozeman, MT 59715
(406) 994-6404
(406) 994-7438 Fax
alex@cns.montana.edu

Travis M. Earles, M.S., M.B.A.
NSTC Representative
Nanobiotechnology
Office of Science and Technology Policy
Executive Office of the President
Room 5203-5
725 17th Street, NW
Washington, DC 20502
(202) 456-6025
(202) 456-6021 Fax
travis_m._earles@ostp.eop.gov

Thomas Earnest, Ph.D.
Senior Scientist/Group Leader
Structural Proteomics Development Group
Physical Biosciences Division
Lawrence Berkeley National Laboratory
Mail Stop 64R0121
1 Cyclotron Road
Berkeley, CA 94720-8118
(510) 486-4603
tnearnest@lbl.gov

Warren J. Ewens, Ph.D.
Professor
Department of Biology
University of Pennsylvania
Leidy Laboratories, Room 324
Philadelphia, PA 19104-6018
(215) 898-7109
(215) 898-8780 Fax
wewens@sas.upenn.edu

Mauro Ferrari, Ph.D., M.S.
Professor
The University of Texas Health Science Center at Houston
Sarofin Research Building, Room 537-D
1825 Pressler Street
Houston, TX 77030
(713) 500-2444
(713) 500-2462 Fax
mauro.ferrari@uth.tmc.edu

Robert A. Gatenby, M.D.
Division Chief
Radiology and Integrated Mathematical Oncology
Moffitt Cancer Center and Research Institute
12902 Magnolia Drive
Tampa, FL 33612
(813) 745-2843
(813) 745-6070 Fax
robert.gatenby@moffitt.org

Daniela S. Gerhard, Ph.D.
Director
Office of Cancer Genomics
National Cancer Institute
National Institutes of Health
Building 31, Room 10A-07
MSC 2580
31 Center Drive
Bethesda, MD 20892-2580
(301) 451-8027
(301) 480-4368 Fax
gerhardd@mail.nih.gov

Robert H. Getzenberg, Ph.D.
Director, Urological Research
Johns Hopkins School of Medicine
Marburg Building, Room 121
600 North Wolfe Street
Baltimore, MD 21287
(410) 502-3137
(410) 502-9336 Fax
rgetzen1@jhmi.edu

Piotr Grodzinski, Ph.D.
Program Director for Cancer Nanotechnology
Office of Technology and Industrial Relations
National Cancer Institute
National Institutes of Health
Building 31
31 Center Drive
Bethesda, MD 20892
(301) 496-1550
(301) 496-7807 Fax
grodzinp@mail.nih.gov

Jordan U. Gutterman, M.D.
Professor of Medicine
Chief
Section of Cellular and Molecular Growth Regulation
The University of Texas M.D. Anderson Cancer Center
Unit 950
1515 Holcombe Boulevard
P.O. Box 301429
Houston, TX 77030-1429
(713) 563-4213
(713) 563-4205 Fax
jgutterm@mdanderson.org

David Haussler, Ph.D., M.S.
Investigator
Howard Hughes Medical Institute
Director
Center for Biomolecular Science and Engineering
University of California, Santa Cruz
Engineering 2 Building, Suite 501
Mail Stop CBS/ITI
1156 High Street
Santa Cruz, CA 95064
(831) 227-2116
(831) 459-1809 Fax
haussler@soe.ucsc.edu

James R. Heath, Ph.D.
Elizabeth W. Gilloon Professor and Professor of Chemistry
Division of Chemistry and Chemical Engineering
California Institute of Technology
Mail Code 127-72
1200 East California Boulevard
Pasadena, CA 91125
(626) 395-6079
(626) 395-2355 Fax
heath@caltech.edu

Henry H.Q. Heng, Ph.D.
Associate Professor
Center for Molecular Medicine and Genetics
School of Medicine
Wayne State University
3226 Scott Hall
540 East Canfield
Detroit, MI 48202
(313) 577-9544
(313) 577-5218 Fax
hheng@med.wayne.edu

W. Daniel Hillis, Ph.D., M.S.
Chairman and Cofounder
Applied Minds, Inc.
1209 Grand Central Avenue
Glendale, CA 91201
(818) 545-1401
(818) 244-0204 Fax
danny@appliedminds.com

Srinivasan S. Iyengar, Ph.D.
Assistant Professor
Chemistry Department
Adjunct Assistant Professor
Physics Department
Indiana University
Room C202B
800 East Kirkwood Avenue
Bloomington, IN 47405-7102
(812) 856-1875
(812) 855-8300 Fax
iyengar@indiana.edu

Paul Janmey, Ph.D.
Professor
Departments of Physiology, Physics, and Bioengineering
School of Medicine
University of Pennsylvania
1010 Vagelos Laboratory
3340 Smith Walk
Philadelphia, PA 19104
(215) 573-7380
(215) 573-6815 Fax
janmey@mail.med.upenn.edu

Don H. Johnson, Ph.D.
J.S. Abercrombie Professor
Department of Electrical and Computer Engineering
Rice University
6100 Main Street
Houston, TX 77005-1892
(713) 348-4956
(713) 348-5685 Fax
dhj@rice.edu

Susan M. Keating, Ph.D.
Senior Scientist
CCS Associates
2005 Landings Drive
Mountain View, CA 94043
(650) 691-4400
(650) 691-4410 Fax
skeating@ccsainc.com

Gary J. Kelloff, M.D.
Special Advisor
Cancer Imaging Program
Division of Cancer Treatment and Diagnosis
National Cancer Institute
National Institutes of Health
Executive Plaza North, Suite 6058
MSC 4910
6130 Executive Boulevard
Bethesda, MD 20892
(301) 594-0423
(301) 480-3507 Fax
kelloffg@mail.nih.gov

Peter Kuhn, Ph.D.
Associate Professor
The Scripps Research Institute
GAC-1200
10550 North Torrey Pines Road
La Jolla, CA 92037
(858) 784-7078
(858) 784-8996 Fax
pkuhn@scripps.edu

Jan Lammerding, Ph.D.
Instructor in Medicine and Associate Biophysicist
Brigham and Women's Hospital/Harvard Medical School
Partners Research Building, Room 283
65 Landsdowne Street
Cambridge, MA 02139
(617) 768-8273
(617) 768-8280 Fax
jlammerding@rics.bwh.harvard.edu

Aurel A. Lazar, Ph.D.
Professor
Department of Electrical Engineering
Columbia University
524 SW Mudd
Mail Code 4705
500 West 120th Street
New York, NY 10027
(212) 854-6438
aurel@ee.columbia.edu

Philip R. LeDuc, Ph.D.
Associate Professor
Departments of Mechanical and Biomedical Engineering and of Biological Sciences
Carnegie Mellon University
Scaife Hall, Room 415
5000 Forbes Avenue
Pittsburgh, PA 15213
(412) 268-2504
(412) 268-3348 Fax
prleduc@cmu.edu

Jerry S.H. Lee, Ph.D.
Special Assistant to the Deputy Director
Office of the Director
National Cancer Institute
National Institutes of Health
Building 31, Room 11A-30C
MSC 2580
31 Center Drive
Bethesda, MD 20892-2580
(301) 594-0255
(301) 496-7807 Fax
leejerry@mail.nih.gov

Rong Li, Ph.D.
Investigator and Professor
Department of Molecular and Integrative Physiology
School of Medicine
University of Kansas
Stowers Institute for Medical Research
1000 East 50th Street
Kansas City, MO 64110
(816) 926-4340
(816) 926-4660 Fax
rli@stowers-institute.org

Jonathan D. Licht, M.D.
Professor and Chief
Division of Hematology/Oncology
Associate Director
Clinical Sciences Research
Robert H. Lurie Comprehensive Cancer Center
Lurie 5-123
303 East Superior Street
Chicago, IL 60611
(312) 503-1114
(312) 503-0189 Fax
j-licht@northwestern.edu

Jan T. Liphardt, Ph.D.
Assistant Professor
Physics Department
Lawrence Berkeley National Laboratory
Stanley Hall, Room 478
Berkeley, CA 94720-3220
(510) 666-2784
liphardt@berkeley.edu

Jennifer Lippincott-Schwartz, Ph.D., M.S.
Chief
Section on Organelle Biology
Cell Biology and Metabolism Branch
National Institute of Child Health and Human Development
National Institutes of Health
Building 18T, Room 101
18 Library Drive
Bethesda, MD 20892
(301) 402-1010
(301) 402-0078 Fax
jlippin@helix.nih.gov

Carlo C. Maley, Ph.D.
Assistant Professor
Molecular and Cellular Oncogenesis Program
Systems Biology Division
The Wistar Institute
3601 Spruce Street
Philadelphia, PA 19104
(215) 495-6838
(215) 495-6829 Fax
cmaley@alum.mit.edu

Scott R. Manalis, Ph.D.
Professor
Massachusetts Institute of Technology
8 Magazine Court
Cambridge, MA 02139
(617) 253-5039
scottm@media.mit.edu

Wallace F. Marshall, Ph.D.
Assistant Professor
Department of Biochemistry and Biophysics
University of California, San Francisco
GH-N376 Genentech Hall
600 16th Street
San Francisco, CA 94143-2200
(415) 514-4323
wmarshall@biochem.ucsf.edu

Owen J.T. McCarty, Ph.D.
Assistant Professor
Department of Biomedical Engineering
Department of Cell and Development Biology
Center for Health and Healing
Oregon Health & Science University
Room 13033
3303 SW Bond Avenue
Beaverton, OR 97006
(503) 418-9307
(503) 418-9311 Fax
mccartyo@ohsu.edu

Lisa Joy McCawley, Ph.D.
Research Assistant Professor
Department of Cancer Biology
Vanderbilt University Medical Center
Preston Building, Room 771
2220 Pierce Avenue
Nashville, TN 37232
(615) 343-9143
(615) 936-2911 Fax
lisa.mccawley@vanderbilt.edu

Robert M. McMeeking, Ph.D.
Professor
Department of Mechanical Engineering
University of California, Santa Barbara
Santa Barbara, CA 93106
(805) 893-8434
(805) 893-8651 Fax
rmcm@engineering.ucsb.edu

Robert Mittman, M.S., M.P.P.
Founder/President
Facilitation, Foresight, Strategy
3 Roberts Court
Moraga, CA 94556
(925) 377-8838
(925) 377-8808 Fax
robert@mittman.org

Raj Mohanty, Ph.D.
Professor
Department of Physics
Boston University
590 Commonwealth Avenue
Boston, MA 02215
(617) 353-9297
(617) 353-9393 Fax
mohanty@bu.edu

Larry A. Nagahara, Ph.D.
Nanotechnology Project Manager
NCI Alliance for Nanotechnology in Cancer
Office of Technology and Industrial Relations
National Cancer Institute
National Institutes of Health
Building 31, Room 10A-52
MSC 2580
31 Center Drive
Bethesda, MD 20892-2580
(301) 496-1550
(301) 496-7807 Fax
nagaharl@mail.nih.gov

Ilya Nemenman, Ph.D.
Research and Development Scientist 4
Computer Computation and Statistical Sciences Division
Los Alamos National Laboratory
Mail Stop B256
Los Alamos, NM 87545
(505) 665-8250
(505) 667-1126 Fax
nemenman@lanl.gov

John E. Niederhuber, M.D.
Director
Office of the Director
National Cancer Institute
National Institutes of Health
Building 31, Room 11A-48
MSC 2590
31 Center Drive
Bethesda, MD 20892-2590
(301) 594-6369
(301) 402-0338 Fax
niederhj@mail.nih.gov

Larry Norton, M.D.
Deputy Physician-in-Chief for Breast Cancer Programs
Memorial Sloan-Kettering Cancer Center
205 East 64th Street
New York, NY 10065
(212) 639-5319
(212) 303-9120 Fax
nortonl@mskcc.org

Thomas V. O'Halloran, Ph.D., M.A.
Charles E. and Emma H. Morrison Professor
Department of Chemistry
Northwestern University
2145 Sheridan Road
Evanston, IL 60208-3113
(847) 491-5060
(847) 491-7713 Fax
t-ohalloran@northwestern.edu

Michael E. Paulaitis, Ph.D.
Ohio Eminent Scholar Professor
Director
Institute in Multiscale Modeling of Biological Interactions
Department of Chemical and Biomolecular Engineering
The Ohio State University
125 Koffolt Laboratories
140 West 19th Avenue
Columbus, OH 43210-1185
(614) 247-8847
(614) 292-3769 Fax
paulaitis.1@osu.edu

Robert Phillips, Ph.D.
Professor of Applied Physics and Mechanical Engineering
Division of Engineering and Applied Science
California Institute of Technology
159 Broad
MC 128-95
1200 California Boulevard
Pasadena, CA 91125
(626) 395-3374
phillips@pboc.caltech.edu

Steven Piantadosi, M.D., Ph.D.
Director
Samuel Oschin Comprehensive Cancer Institute
Cedars-Sinai Medical Center
Mezzanine, Room C2002
8700 Beverly Boulevard
Los Angeles, CA 90048
(310) 423-8431
(310) 423-8300 Fax
steven.piantadosi@csmc.edu

Hong Qian, Ph.D.
Professor
Department of Applied Mathematics
University of Washington
Box 352420
Seattle, WA 98195
(206) 543-2584
(206) 685-1440 Fax
qian@amath.washington.edu

Brian J. Reid, M.D., Ph.D.
Full Member
Divisions of Human Biology and Public Health Sciences
Fred Hutchinson Cancer Research Center
Mail Stop C1-157
1100 Fairview Avenue, North
Seattle, WA 98109-1024
(206) 667-4073
(206) 667-6132 Fax
bjr@fhcrc.org

Cynthia A. Reinhart-King, Ph.D.
Assistant Professor
Department of Biomedical Engineering
Cornell University
Weill Hall, Room 302
Ithaca, NY 14850
(607) 255-8491
cak57@cornell.edu

Joseph A. Rudnick, Ph.D.
Acting Dean
Professor of Physics
Division of Physical Sciences
College of Letters and Science
University of California, Los Angeles
Box 951438
Murphy Hall, Room 2300
Los Angeles, CA 90095-1438
(310) 825-1042
(310) 825-7823 Fax
jrudnick@physics.ucla.edu

Joel H. Saltz, M.D., Ph.D.
Chief Medical Information Officer
Director
Comprehensive Center for Informatics
Professor
Department of Pathology
Emory University
Suite 500
1521 Dickey Drive
Atlanta, GA 30322
(404) 727-6202
(404) 727-4992 Fax
jhsaltz@emory.edu

Thomas D. Schneider, Ph.D.
Research Scientist
National Cancer Institute
National Institutes of Health
Building 469, Room 105
Frederick, MD 21702-1201
(301) 846-5581
(301) 846-5598 Fax
toms@ncifcrf.gov

James P. Sethna, Ph.D., M.A.
Professor of Physics
Cornell University
Clark Hall
Ithaca, NY 14853-2501
(607) 255-5132
(607) 255-6428 Fax
sethna@lassp.cornell.edu

Phillip A. Sharp, Ph.D.
Professor
The David H. Koch Institute for Integrative Cancer Research
Massachusetts Institute of Technology
Room E17-529
77 Massachusetts Avenue
Cambridge, MA 02139-4307
(617) 253-6421
(617) 253-3867 Fax
sharppa@mit.edu

Darryl K. Shibata, M.D.
Professor of Pathology
Keck School of Medicine
University of Southern California
Norris Cancer Center, Room 6410
1441 Eastlake Avenue
Los Angeles, CA 90033
(323) 226-7067
(323) 226-2686 Fax
dshibata@usc.edu

James L. Siegrist, Ph.D.
General Sciences Associate Director
Director
Physics Division
Lawrence Berkeley National Laboratory
Mail Stop 50R4049
1 Cyclotron Road
Berkeley, CA 94720
(510) 486-4397
(510) 486-6003 Fax
jlsiegrist@lbl.gov

Caroline C. Sigman, Ph.D.
President
CCS Associates
2005 Landings Drive
Mountain View, CA 94043
(650) 691-4400
(650) 240-4013 Fax
csigman@ccsainc.com

Dinah S. Singer, Ph.D.
Director
Division of Cancer Biology
National Cancer Institute
National Institutes of Health
Executive Plaza North, Suite 5044
6130 Executive Boulevard
Bethesda, MD 20892-7390
(301) 496-8636
(301) 496-8656 Fax
ds13j@nih.gov

Ana M. Soto, M.D.
Professor
Department of Anatomy and Cellular Biology
Tufts University
Arnold Building, Room 116 Harrison
Boston, MA
(617) 636-6954
(617) 636-6536 Fax
ana.soto@tufts.edu

Fuyuhiko Tamanoi, Ph.D.
Professor
Department of Microbiology, Immunology, and Molecular Genetics
California NanoSystems Institute
University of California, Los Angeles
Molecular Science Building, Room 1602
609 Charles E. Young Drive
Los Angeles, CA 90095-1489
(310) 206-7318
(310) 206-5231 Fax
fuyut@microbio.ucla.edu

Peter J. Thomas, Ph.D.
Associate Professor
Department of Mathematics, Biology, and Cognitive Science
Case Western Reserve University
Yost 210
Cleveland, OH 44106
(216) 368-3623
peter.j.thomas@case.edu

Thomas G. Thundat, Ph.D.
Distinguished Scientist
Oak Ridge National Laboratory
Building 4500S
MS 6123
Oak Ridge, TN 37831-6123
(865) 574-6201
(865) 574-6210 Fax
ugt@ornl.gov

Yiider Tseng, Ph.D.
Associate Professor
Department of Chemical Engineering
University of Florida
Chemical Engineering Building, Room 223
Museum Road
Gainesville, FL 32611
(352) 392-0862
(352) 392-9513 Fax
ytseng@che.ufl.edu

Yu-Li Wang, Ph.D.
Professor and Head
Department of Biomedical Engineering
Carnegie Mellon University
Suite 4105-07
700 Technology Drive
Pittsburgh, PA 15219
(412) 268-4442
(412) 268-9807 Fax
yuliwanga@andrew.cmu.edu

Samuel A. Wells, Jr., M.D.
Professor
Department of Surgery
Washington University School of Medicine
Campus Box 8109
660 South Euclid Avenue
St. Louis, MO 63110
(919) 201-0310
(314) 454-1898
wellss@wudosis.wustl.edu

Robert M. Westervelt, Ph.D.
Professor
Department of Physics
School of Engineering and Applied Sciences
Harvard University
29 Oxford Street
Cambridge, MA 02138
(617) 495-3296
(617) 495-9837 Fax
westervelt@seas.harvard.edu

John P. Wikswo, Ph.D.
Gordon A. Cain University Professor
Department of Physics and Astronomy
Vanderbilt University
VU Station B 351807
Nashville, TN 37235-1807
(615) 343-4124
(615) 322-4977 Fax
john.wikswo@vanderbilt.edu

Denis Wirtz, Ph.D.
Professor
Department of Chemical and Biomolecular Engineering
Johns Hopkins University
Maryland Hall, Room 116
3400 North Charles Street
Baltimore, MD 21218
(410) 516-7006
(410) 516-2355 Fax
wirtz@jhu.edu

Miqin Zhang, Ph.D.
Associate Professor
Department of Materials Science and Engineering
University of Washington
302L Roberts Hall
Box 352120
Seattle, WA 98195
(206) 616-9356
(206) 543-3100 Fax
mzhang@u.washington.edu

NATIONAL® CANCER INSTITUTE

NIH Publication No. 09-7388
Printed June 2009