

## Meeting Report

## Systems Biology: Confronting the Complexity of Cancer

Andrew J. Gentles<sup>1</sup> and Daniel Gallahan<sup>2</sup>

## Abstract

The AACR-NCI Conference "Systems Biology: Confronting the Complexity of Cancer" took place from February 27 to March 2, 2011, in San Diego, CA. Several themes resonated during the meeting, notably (i) the need for better methods to distill insights from large-scale networks, (ii) the importance of integrating multiple data types in constructing more realistic models, (iii) challenges in translating insights about tumorigenic mechanisms into therapeutic interventions, and (iv) the role of the tumor microenvironment, at the physical, cellular, and molecular levels. The meeting highlighted concrete applications of systems biology to cancer, and the value of collaboration between interdisciplinary researchers in attacking formidable problems. *Cancer Res*; 71(18); 5961–4. ©2011 AACR.

## Introduction

Our knowledge of cancer is advancing at an unprecedented rate; however, the more we understand, the more complex this family of diseases reveals itself to be. Various processes, at both a molecular and cellular level, function in a dynamic and connective manner which must be understood to address the disease. These processes are guided by the genetic and epigenetic events inherent in cancer cells, and in the microenvironment with which they interact. The rate of innovation in technologies for interrogating cellular systems at multiple levels presents major challenges in interpretation and in uncovering core insights. Systems biology has developed to address such issues by bringing to bear the techniques of computational and mathematical sciences in an attempt to tame some of the complexity. The National Cancer Institute (NCI)'s Integrative Cancer Biology Program was among the earliest to explicitly promote multidisciplinary approaches for advancing cancer research. Recently, the AACR and NCI held their first joint conference, "Systems Biology: Addressing the Complexity of Cancer," presenting an opportunity to review the current state of the field and highlight ongoing challenges. The meeting covered a vast spectrum of topics, befitting the subject. We outline some of the results presented, with apologies to those researchers whose work cannot be included due to space.

## Genetic Systems

The spectrum of mutational changes differs widely between cancers, both in terms of their overall frequency, and the relative frequency of mutational classes such as transitions versus transversions. Recent results from The Cancer Genome Atlas (TCGA) presented by Paul Spellman (Oregon Health Science University, Portland, OR) and Gad Getz (Broad Institute, Boston, MA) show that neuroblastomas, prostate, and breast cancers exhibit small numbers of mutations (~10 per exome, or 0.3/Mb, in pediatric cancers), whereas lung cancer and UV-induced melanoma (~3,000 per exome) have orders of magnitude more. This likely reflects the role of carcinogens in the latter malignancies. Moreover, the specific classes of mutations observed differ markedly in melanoma (UV-induced) and lung cancer (smoking-related; Getz and colleagues, in preparation). Although 96% of ovarian tumors harbor P53 mutations, recurrent mutations are otherwise scarce, with many genes being mutated in only a few percent of cases. However, particular genes (e.g., *IDH1* and *MYD88*) are commonly amplified or mutated, and implicated across malignancies. At a higher level, glioblastomas show few structural genomic aberrations, whereas ovarian cancers exhibit extensive genome "shredding." The reasons for these differences are unknown.

## Molecular Systems

The genetic configuration is integrated through intricate signaling pathways culminating in the activation or repression of cellular process that are indicative of, or contribute to, tumor initiation and progression. One common activity in systems biology is to reconstruct large interaction networks, at intracellular and intercellular levels. However, replacing a list of differentially expressed genes with a massive hairball is not terribly elucidating. The key is identifying dysregulation of specific genes, interactions,

**Authors' Affiliations:** <sup>1</sup>Center for Cancer Systems Biology and Department of Radiology, School of Medicine, Stanford University, Stanford, California; and <sup>2</sup>National Cancer Institute, Bethesda, Maryland

**Corresponding Author:** Andrew J. Gentles, Stanford University, P060 Lucas Center, Stanford, CA 94305. Phone: 650-725-3121; Fax: 650-723-5795; E-mail: andrewg@stanford.edu or Daniel Gallahan. E-mail: gallahad@mail.nih.gov

doi: 10.1158/0008-5472.CAN-11-1569

©2011 American Association for Cancer Research.

and subnetworks that drive malignancy, influence phenotypes, and characterize response to therapy. Both top-down and bottom-up approaches have important roles to play.

Lou Staudt (National Cancer Institute, Bethesda, MD) described RNAi screening of activated B-cell type diffuse large B-cell lymphoma (ABC DLBCL) cell lines, which identified the cytosolic adaptor protein gene *MYD88* as essential to their survival (1). High throughput resequencing showed the gene to be mutated and constitutively activated in the cell lines, and 29% of clinical ABC DLBCL samples. Similarly, *JAK2* (Janus kinase 2) and *JMJD2C* (a Jumonji domain histone demethylase) were found to cooperatively remodel the epigenome in primary mediastinal B-cell lymphoma and Hodgkins lymphoma (2). Intriguingly, despite heterogeneity in oncogenic mechanisms, cancer cells often seem to "hardwire" prosurvival/proliferative pathways already utilized by their normal cell of origin.

A small but remarkable number of neuroblastoma patients, even with massive metastatic burden, experience spontaneous differentiation or complete regression of their disease. John Maris (Children's Hospital of Pennsylvania, Philadelphia, PA) described a top-down approach to analyzing neuroblastoma by using a Master Regulator Analysis method (MRA, described below) developed by Andrea Califano's group at Columbia University (New York, NY). By comparing patients with good or poor prognosis, *ALK* was identified as a key driver gene in familial neuroblastoma. Sixteen percent of cases subjected to high-throughput resequencing were found to harbor mutations in *ALK*, consistent with previous reports (3).

DNA repair pathways defects are present in the majority of cancers. The Mre11-Rad50-Nbs1 (MRN) trimeric complex can sense diverse lesion types and coordinate appropriate DNA damage responses. John Tainer (Lawrence Berkeley National Laboratory, Berkeley, CA) described how this ability is attributable to the multiple possible conformational states of its subunits (4). MRN functions as a molecular machine, altering shape according to the type of DNA damage present, and simultaneously integrating correct cellular response mechanisms. This dynamic process highlights another layer of molecular complexity which may be therapeutically exploitable by employing conformational inhibitors.

#### Network regulation of cell function and dysfunction

Even as scientists, we have a tendency for presenting results in a "Just So" fashion. The very concept of a pathway is nebulous, given tissue and context-specific differences in the information flow through cellular networks. Top-down data-driven network reconstruction approaches combat this need for story telling but require better downstream methods for making specific inferences about causal associations with disease.

The MRA approach described by Andrea Califano identifies highly connected regulatory "hubs" using the ARACNe method (5). Targets of these hub genes are then used in gene set analysis. Coherent changes in predicted targets between, for example, 2 cancer subtypes "point back" to the driver gene. Application to differences between glioblastomas with mesenchymal versus proneural expression signatures identified a core network involving CEPBP $\alpha$  and STAT3 that was validated to synergistically drive the mesenchymal phenotype

(6). Causality is most readily inferred in the presence of perturbations; however, most (e.g., siRNA) are artificial. DNA variations such as SNPs can be considered as natural perturbations for informing the task of causal inference. Ernest Fraenkel (MIT, Boston, MA) described 1 method for integrating such diverse data types using Prize-collecting Steiner trees (7). Given an interactome, this approach balances the "cost" of removing nodes from the network against the cost of using edges to include nodes. Penalties depend on the reliability of edges in the network, and strength of support of nodes in the experimental data. Integration of expression data, phosphoprotein data, and DNAase sensitivity sites implicates a central role for ESRI in glioblastoma.

David Haussler (University of California, Santa Cruz, CA) outlined how TCGA data is integrated into the UCSC Cancer Genomics Browser. A key modeling component is the PARADIGM method developed with Josh Stuart's group (University of California, Santa Cruz, CA). This system uses a "central dogma" based model coupled with Bayesian integration of multiple data types to infer biological processes that are activated or repressed in a patient-specific manner (8). Initial analysis identified a previously unknown key role for a FOXM1-driven network in ovarian cancer. Paul Spellman (standing in for Joe Gray, Oregon Health Science University, Portland, OR) outlined work on assaying phenotypic responses of 70 breast cancer cell lines to a large number of drugs under many conditions (9). Genomic data including transcriptome and exome sequencing, CNVs, fusion transcripts, alternative splicing events, and protein signaling are integrated using PARADIGM to identify subtype-specific pathway dysregulation. Subsequent mapping to patient tumor data makes selection of specific drug interventions possible.

Joel Wagner (Doug Lauffenburger lab, MIT, Boston, MA) applied Bayesian networks to reconstruct dynamic signaling networks downstream of surface receptor tyrosine kinase (RTK). Six diverse RTKs were transfected into a cell line after which protein lysate microarrays were used to assay phosphorylation and PARP cleavage at 11 time points, in the presence of 91 shRNA perturbations. Time-invariant edges in the network were conserved across multiple receptors, whereas time-variant edges were conserved in fewer. This supports a "backbone" of signaling across RTKs, layered on which are RTK-specific signaling features affecting phenotype and signal timing. Jason Neil (Forest White lab, MIT, Boston, MA) presented a network model of how EMT-mediated rewiring of epidermal growth factor receptor (EGFR) signaling influences cell mobility. iTRAQ and LC-MS data showed dynamic changes in pre- and post-EMT phosphorylation of SRC and ERK kinases downstream of EGFR. Further characterization may present targets for modulating EGF-driven cell motility.

The static connectivity of networks in different conditions may be less informative than differences between them. Trey Ideker (University of California, San Diego, CA) presented work in yeast where genetic interaction maps were generated under normal and methyl methanesulfonate-induced DNA damage conditions using epistatic miniarrays. DNA damage-response genes were no more likely than random to exhibit association with each other in either condition

considered separately (10). However, interactions between these genes were strongly enriched in the dynamic DNA damage map representing differences between the networks.

### Cellular systems

Cancer represents a heterogeneous collection of cell types and differentiation states that function as an "organ." Cancer stem cells might play a critical role in the generation and maintenance of this heterogeneous environment (11), including the tumor, stroma, and also systemic host cellular systems such as the immune system. One may also think of the host's microbiome or bacterial population as an interacting system, because 90% of cells in the human body are actually microbial with approximately 400 phylotypes inhabiting the typical gut. Elhanan Borenstein (University of Washington, Seattle, WA) discussed reconstruction of cross-phylum networks using shotgun metagenomics to look at total sets of genes and follow their interactions. These, and related computational techniques, may prove useful in probing and understanding tumor-stromal interactions in cancer (12).

### Cancer stem cells and the tumor microenvironment

Induced pluripotent stem (iPS) cells become more similar to ES cells in gene expression and epigenetic patterns with increasing numbers of passage in culture. Kathrin Plath (University of California, Los Angeles, CA) outlined work on pre-iPS cells to understand late stages of reprogramming (13). Many genes that are cobound by the reprogramming factors *Oct4*, *Sox*, and *Klf4* (OSK) in fully reprogrammed iPS cells and ES cells are not cobound in pre-iPS cells; however, the cMYC network becomes engaged significantly earlier than the OSK network. There are tempting parallels to be drawn between iPS reprogramming and the possible reactivation or subversion of self-renewal functions in cancer cells, and this is likely to be an informative direction of research.

If subpopulations of tumor initiating cells (TIC) maintain a tumor, eliminating them is necessary for effective treatment. Stephen Wong (Methodist University, Houston, TX) posed the hypothesis that TICs are more likely to be high drug efflux cells. His group conducted a screen of 1,280 compounds, coupled with automated image acquisition and processing to identify and isolate high efflux cells in malignant populations (14). Several inhibitors of drug efflux found in the screen sensitized lung cancer cells to chemotherapeutics, overcoming multiple-drug resistance. The software tools deployed in this approach are freely available to the community.

Cancer has been eradicated thousands of times in a Petri dish, but these results rarely translate to *in vivo*. Tumors are dynamic and evolving ecosystems of interacting cancer cells, endothelial cells, fibroblasts, and infiltrating immune cells, among others. Moreover, the physical properties of the tumor microenvironment modulate cellular behavior in poorly understood ways.

Collateral damage of the host immune system is another factor in treatment failure, not just due to impairment of its antitumor functions. Peter Nelson (Fred Hutchinson Cancer Research Center, Seattle, WA) showed that introducing DNA damage to fibroblasts, as might occur during chemotherapy,

dramatically affects their gene expression profile, upregulating WNT16B 60-fold at the transcript level, with increased protein excretion. This paracrine signal promotes an aggressive mesenchymal phenotype in tumor cells. Valerie Weaver (University of California, San Francisco, CA) described her lab's work on the mechanical effect of the extracellular matrix (ECM) on tumor cells. ECM stiffness promoted malignant progression, and regulated sensitivity to chemotherapy and radiation (15). As the medium is changed from soft to stiff a "switch" occurs in gene expression profiles, with associated changes in chromatin structure and genomic methylation patterns. A specific SMRT/NCOR2 signature affects chemotherapy response, and similar stress pathway agonists are also potential modulators of tumor behavior.

"Triple-negative" breast cancer (TNBC) patients have dismal prognosis with no specific effective treatments. Jennifer Pietenpol (Vanderbilt University, Nashville, TN) discussed meta-analysis of 21 gene expression data sets comprising 587 TNBCs. Seven subtypes could be discerned related to cell cycle, DNA damage response, immune response, TGF $\beta$ /mesenchymal signatures, focal adhesion, and androgen signaling. Cell lines representing these subtypes were designated according to gene expression profiles, and potential regulatory signaling pathways targeted pharmacologically.

### System approaches to drug discovery

Given knowledge of tumorigenic mechanisms, ideally at a patient-specific level, the next challenge for clinical translation is to find therapeutic interventions that can target them. Cell line and *in vivo* models provide initial guidance, but dosage and treatment timing are crucial, particularly in the context of combination regimens.

The Connectivity map (C-map) is one tool for initial identification of potential drug interventions (16). Given a gene expression signature, the aim is to identify drugs which induce an opposite pattern. Todd Golub (Broad Institute, Boston, MA) described the future evolution of the C-map to encompass 5,000 compounds, 3,000 RNAi perturbations, and 3,000 over-expression experiments in 20 cell lines at multiple time points and doses. A total of 1,000 representative "landmark genes" have been identified from which expression levels of other genes can be imputed with 80% accuracy, permitting the use of a bead-based Luminex system, dramatically reducing costs.

In another demonstration of the power of RNAi, Michael Hemann (MIT, Boston, MA) described an *in vivo* mouse (importantly, with an intact immune system) screen to identify predictors of response to drug interventions. Libraries of shRNAs were used to infect lymphoma/leukemia cells which were transplanted *in vivo* to see which sensitized tumors to chemotherapy. Dependence of lymphoma and leukemia cells on members of the BCL-2 family was readily identified. By clustering drugs by genotype sensitivity, an 8 shRNA signature was found that could identify genes associated with response of mammalian cells to genotoxic drugs (17).

Even if dosing levels and schedule are established for a combination regimen, ensuring patient adherence is challenging. Yet as illustrated by Michael Lee (Michael Yaffe lab, MIT, Boston, MA), timing and ordering of treatments is critical to

their synergy. In a drug screen TNBCs, staggered treatment with erlotinib followed by doxylamine had a much greater effect than doxylamine followed by erlotinib. Chronic EGFR inhibition by erlotinib rewires signaling in BT20 cell lines, sensitizing them to DNA damage. A partial least squares model based on a 35-protein phosphoprotein assay indicated that EGFR inhibition reactivates a Caspase-8 dependent apoptotic pathway required for DNA damage-induced killing. Activity of the pathway is normally repressed in BT20, rendering them insensitive to doxylamine.

## Summary

The prognosis for many cancer patients remains dismal. Countless treatments fail in late stage clinical trials despite promising early results. How many might have been successful with a better comprehension of the importance of timing and ordering of treatments? The role of the immune system and microenvironment has deservedly received increasing attention—now including physical properties, not just biological ones. It is clear that some treatments fail not just by damaging the immune system but by causing it to backfire and promote tumor progression.

The observed mutational frequency across cancers differs. How many rare mutations, which are difficult to detect even if sequencing were 99.9999% accurate, are actually drivers or at least copilots of oncogenic phenotypes? Another challenge is to study the continuum and evolution of cancer—from early initiation events through progression and treatment response. Of 500 ovarian tumors acquired by TCGA, only 5 represent stage 2 disease, and none represent the ovary-confined stage 1, due to availability of suitable quality samples. Yet analyzing and understanding early stage disease is absolutely critical.

## References

1. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature* 2011;470:115–9.
2. Rui L, Emre N, Kruhlak MJ, Chung HJ, Steidl C, Slack G, et al. Cooperative Epigenetic Modulation by Cancer Amplicon Genes. *Cancer Cell* 2010;18:590–605.
3. Mossé YP, Laudenslager M, Longo L, Cole KA, Wood A, Attiyeh EF, et al. Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature* 2008;455:930–5.
4. Williams GJ, Williams RS, Williams JS, Moncalian G, Arvai AS, Limbo O, et al. ABC ATPase signature helices in Rad50 link nucleotide state to Mre11 interface for DNA repair. *Nat Struct Mol Biol* 2011;18:423–31.
5. Margolin A, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Favera R, et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 2006;7 (Suppl 1):S7.
6. Carro MS, Lim WK, Alvarez MJ, Bollo RJ, Zhao X, Snyder EY, et al. The transcriptional network for mesenchymal transformation of brain tumours. *Nature* 2009;463:318–25.
7. Huang S, Fraenkel E. Integrating proteomic, transcriptional, and interactome data reveals hidden components of signaling and regulatory networks. *Sci Signal* 2009;2:ra40.
8. Vaske CJ, Benz SC, Sanborn JZ, Earl D, Szeto C, Zhu J, et al. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* 2010;26:i237.
9. Heiser LM, Wang NJ, Talcott CL, Laderoute KR, Knapp M, Guan Y, et al. Integrated analysis of breast cancer cell lines reveals unique signaling pathways. *Genome Biol* 2009;10:R31.
10. Bandyopadhyay S, Mehta M, Kuo D, Sung MK, Chuang R, Jaehnig EJ, et al. Rewiring of genetic networks in response to DNA damage. *Science* 2010;330:1385.
11. Chaffer CL, Brueckmann I, Scheel C, Kaestli AJ, Wiggins PA, Rodrigues LO, et al. Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc Natl Acad Sci U S A* 2011;108:7950–5.
12. Borenstein E, Kupiec M, Feldman MW, Ruppin E. Large-scale reconstruction and phylogenetic analysis of metabolic environments. *Proc Natl Acad Sci U S A* 2008;105:14482–7.
13. Plath K, Lowry WE. Progress in understanding reprogramming to the induced pluripotent state. *Nat Rev Genet* 2011;12:253–65.
14. Xia X, Yang J, Li F, Li Y, Zhou X, Dai Y, et al. Image-Based Chemical Screening Identifies Drug Efflux Inhibitors in Lung Cancer Cells. *Cancer Res* 2010;70:7723.
15. Butcher DT, Alliston T, Weaver VM. A tense situation: forcing tumour progression. *Nat Rev Cancer* 2009;9:108–22.
16. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006;313:1929.
17. Jiang H, Pritchard JR, Williams RT, Lauffenburger DA, Hemann MT. A mammalian functional-genetic approach to characterizing cancer therapeutics. *Nat Chem Biol* 2011;7:92–100.

This problem will become more severe as new technologies and modalities are introduced.

We have vastly increased our understanding of cancer, but perhaps the unifying result has been the recognition of the complexity of these diseases. Despite skepticism in some quarters, systems biology has made significant contributions. At this meeting, speakers described advances in understanding oncogenic mechanisms using computational techniques from data-driven top-down models, such as MRA and Bayesian networks, to detailed mechanistic and dynamic models of specific processes. Methods such as PARADIGM are tackling the question of identifying patient-specific oncogenic mechanisms and assisting identification of appropriate therapeutics. The next few years will see acceleration in the rates and types of data acquisition, such as whole genome sequencing. Making sense of this information will necessitate sophisticated models developed in interdisciplinary collaborations, and there will be challenges in balancing patient privacy with the need for making data accessible as widely as possible. Support and understanding from patients and their advocates will be important in guiding these efforts.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Acknowledgments

We thank the cochairs Andrea Califano, Dan Gallahan, Todd Golub, Joe Gray, Doug Lauffenburger, and Jennifer Pietenpol for putting together an excellent schedule and meeting, as well as the AACR and the NCI for their continuing support of systems biology of cancer.

Received May 10, 2011; accepted June 18, 2011; published OnlineFirst September 6, 2011.

# Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

## Systems Biology: Confronting the Complexity of Cancer

Andrew J. Gentles and Daniel Gallahan

*Cancer Res* 2011;71:5961-5964. Published OnlineFirst September 13, 2011.

**Updated version** Access the most recent version of this article at:  
doi:[10.1158/0008-5472.CAN-11-1569](https://doi.org/10.1158/0008-5472.CAN-11-1569)

**Cited articles** This article cites 17 articles, 7 of which you can access for free at:  
<http://cancerres.aacrjournals.org/content/71/18/5961.full#ref-list-1>

**E-mail alerts** [Sign up to receive free email-alerts](#) related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at [pubs@aacr.org](mailto:pubs@aacr.org).

**Permissions** To request permission to re-use all or part of this article, contact the AACR Publications Department at [permissions@aacr.org](mailto:permissions@aacr.org).